Substitute or complement? Defining the relative place of EEG and fMRI in the detection of voluntary brain reactions

Authors: Damien Gabriel[1,2*]; Julie Henriques[3]; Alexandre Comte[1,2,4]; Lyudmila Grigoryeva[3]; Juan-Pablo Ortega[3,5]; Elodie Cretin[2,6]; Gaëlle Brunotte[2]; Emmanuel Haffen[1,2,7]; Thierry Moulin[1,2,4]; Régis Aubry[2,8]; Lionel Pazart[2]

Affiliations:

[1] Laboratoire de Neurosciences de Besançon, Université de Franche-Comté

[2] Centre d'investigation Clinique–Innovation Technologique, Inserm 808, CHRU de Besançon

[3] Laboratoire de Mathématiques de Besançon, Université de Franche-Comté

[4] Service de neurologie, CHRU Besançon

[5] Centre National de la Recherche Scientifique

[6] Espace éthique Bourgogne Franche-Comte

[7] Service de psychiatrie de l'adulte, CHRU de Besançon

[8] Département douleur soins palliatifs, CHRU de Besançon

***Corresponding author:**

Damien Gabriel, Ph.D ; Clinical Investigation Center, Bâtiment Saint-Joseph, CHRU Saint Jacques, 2 place Saint Jacques, F-25 030 Besançon ; Tel: +33 3 81 21 91 48; Fax: +33 3 81 21 90 28; email: dgabriel@chu-besancon.fr

Abstract

To improve the assessment of awareness in patients with disorders of consciousness, recent protocols using fMRI have been developed, and led some specialized coma centers to use this method on a routine basis. Recently, promising results have also been observed with EEG, a less expensive and widely available technique. However, since the spatiotemporal nature of the recorded signal differs between both EEG and fMRI, the question of whether one method could substitute or should complement the other method is a matter of debate. In this study, we compared the neural processes of two well-known EEG and fMRI mental imagery protocols to define the relative place of each method in the assessment of awareness.

A group of 20 healthy volunteers performed both EEG and fMRI command-following and communication tasks. Distinct command following was found with both EEG and fMRI for 5 subjects, only with fMRI for 12 subjects, and only with EEG for 1 subject. In the communication task, neither EEG nor fMRI alone gave satisfactory results and no reliable communication could be established in approximately 1/3rd of the participants.

If fMRI showed the best performance to detect volitional reactions in mental imagery tasks, our results provide evidence that the use of EEG must not be underestimated since a better detection was found with this method for at least one subject. More than being used as a substitute, EEG should complement fMRI to improve the detection of sign of awareness, and to reduce the risks of misjudgments.

Key words: Consciousness, EEG, functional MRI, neuroimaging, mental imagery

1.Introduction

Recent advances in brain imaging have led to the development of new methods of detecting awareness in patients with disorders of consciousness (DOCs), such as patients with an unresponsive wakefulness syndrome (UWS, formerly known as vegetative state) or patients in a minimally conscious state (MCS). In the absence of overt behavioral responses from these patients, imaging-based diagnostic methods can be efficiently used to reveal covert and volitional reactions. Voluntary activation of specific brain regions has been observed repeatedly on functional Magnetic Resonance Imaging (fMRI) in some patients with DOCs, similarly to what is observed in healthy control subjects. Several studies have indeed shown that a subset of patients were able to generate meaningful brain responses when they were asked to guide their attention to specific stimuli (Naci and Owen, 2013) or to perform mental imagery tasks such as playing tennis or moving around a familiar place (Owen et al., 2006). The same technology successfully enabled a rudimentary yes-or-no communication to be established with some patients able to follow commands (Fernandez-Espejo and Owen, 2013; Monti et al., 2010; Naci and Owen, 2013).

Since fMRI investigations are useful to detect inconsistencies in the clinical assessment of patients with DOCs, specialized coma centers in Canada, USA, United Kingdom and Belgium are now using these methods on a routine basis (Fernandez-Espejo and Owen, 2013). However, for various ethical and medical reasons, such as the cost of an fMRI exam, the low availability of a scanner, and the necessity of transferring the patient to the machine, which is not possible in some cases, fMRI assessments of awareness are difficult to generalize to all hospitals. As a consequence, several research groups have begun to focus more on electroencephalography (EEG), which is less expensive and transportable to the patient bedside. With protocols similar to fMRI, some EEG studies showed that patients with DOCs were able to willfully direct their attention towards specific stimuli (Bekinschtein et al., 2009; Chennu et al., 2013; Schnakers et al., 2008) or to perform mental imagery tasks (Cruse et al., 2011; Cruse et al., 2012a; Goldfine et al., 2011; Holler et al., 2013). For example, in the experiment from (Cruse et al., 2011; Cruse et al., 2012a), patients with DOCs were asked to imagine either squeezing their right hand or wiggling their toes. The classification and the comparison between these two tasks, made with a multivariate automatic method (Support Vector Machine), revealed that 3 out 16 patients and 5 out of 22 patients initially diagnosed as being respectively in a UWS and in a MCS were able to willfully perform the tasks as instructed, suggesting that they were actually aware. Although the reliability of these results has been debated (Goldfine et al., 2013) and that this method showed a somewhat reduced sensitivity compared to its equivalent fMRI protocol, it opens the possibility to use EEG mental imagery tasks in routine care for the evaluation of awareness (Fernandez-Espejo and Owen, 2013). Moreover, EEG may also be used to communicate with patients with DOCs, similarly to what has been achieved with fMRI.

It is not clear, however, what the real place of EEG is when compared to fMRI. Should EEG be considered as a substitute for, or as a complement to fMRI? It is well known that the cerebral nature of the signal measured with both methods is not the same. In fMRI, the distinction between two mental imagery tasks is reflected through changes of cerebral blood flow over specific brain areas (Ghaem et al., 1997; Lotze et al., 1999). These changes are best discriminated spatially, which implies that the selected mental imagery tasks have to activate clearly distinct brain areas. In EEG, the differences in cortical activity are best detected in the temporal domain, by variations of the oscillatory activity. Planning a finger movement is thus known to block or desynchronize the mu (8Hz-12Hz) and/or the beta (13Hz-30Hz) rhythm over the motor cortex whereas imagining a foot movement can enhance them (Pfurtscheller et al., 2006; Pfurtscheller and Lopes da Silva, 1999). In that respect, if a few recent EEG studies compared EEG and fMRI command-following in some patients with DOCs, (Chennu et al., 2013; Cruse et al., 2013; Gibson et al., 2014), a problematic issue is how to interpret similarities and discrepancies in the results. In the absence of a full understanding of the neural correlates of awareness, it is extremely complex to disentangle purely methodological differences, a reduced sensitivity of one neuroimaging method for example, from differences specifically related to the pathology. This is all the more important that the potential strategies considered by coma centers wishing to assess awareness in patients with DOCs, depend on this understanding. Do both EEG and fMRI mental imagery protocols give similar results or should one method be preferred to the other, or should both methods be systematically used in conjunction?

In the present study, our goal was to assess the relative place of EEG and fMRI mental imagery protocols in assessing awareness. To that purpose, a group of 20 healthy volunteers was tested with two standard mental imagery tasks, one with fMRI (Monti et al., 2010) and one with EEG (Cruse et al., 2011; Cruse et al., 2012a). The design of the EEG protocol was modified so as to be also used as a communication tool. We investigated whether the EEG mental imagery tasks could be used to detect covert awareness and to communicate as efficiently as with fMRI in aware subjects, whether a subject classified as aware with one neuroimaging method would necessarily be classified as aware with the other method, or whether each method would provide additional information concerning the presence of awareness. The implications of these results on the care of patients with DOCs will then be discussed.

2. Experimental procedures

2.1. Participants

20 neurologically healthy right-handed adults (11 female; 9 male) aged between 25 and 66 (mean: 35.6) participated in this study. The local ethical committee approved this study, and written informed consents were obtained from all participants.

2.2. Experimental procedure

All participants were first tested with the coma recovery scale to ensure that all could be classified as aware. As expected, a perfect score was obtained for all of them. Next, they participated in two mental imagery and communication tasks used to detect covert awareness in patients with DOCs, one with fMRI and one with EEG. The fMRI task was replicated from the study of (Monti et al., 2010). The EEG task was the same as (Cruse et al., 2011; Cruse et al., 2012a), but with an additional communication part. The communication part was composed of 3 autobiographical questions of which the answers were collected prior to the experiment. At the end of both experiments, questionnaires were given to the subjects to evaluate which mental imagery task was considered as the most difficult.

2.2.1. Command-following task

2.2.1.1. EEG

The command-following task was the same as (Cruse et al., 2011). Each participant had to explicitly perform either right-hand imagery or toe imagery each time he or she heard a beep. For the right-hand imagery task, the instruction was to imagine squeezing the right hand into a fist and then to relax it each time the participant would hear a beep. For the toe, the instruction was to imagine wiggling all of the toes on both feet, and then to relax them. Each block began with the auditory presentation of the task instructions, followed by the binaural presentation of 15 tones (600 Hz, 60 ms duration) with an inter-stimulus interval varying randomly between 4.5 and 9.5 seconds. All of the subjects completed a total of 8 blocks of command-following tasks (4 right-hand imagery, and 4 toe imagery) presented in a pseudorandomized order. A break was provided to participants before starting the next block.

2.2.1.2. fMRI

The first two runs were part of the localizer task. In this task, participants were explicitly instructed to perform two types of mental imagery: motor imagery and spatial navigation. In each run and every 30 seconds, an auditory cue word marked the beginning of a given period ("tennis" or "visit" for mental imagery periods and "relax" for rest periods). In the motor imagery task, participants were instructed to imagine hitting a tennis ball by moving their arm back and forth each time they heard the spoken cue word "tennis". They were told only to move their right arm back and forth to hit the ball and not to imagine the court or to visualize any opponent. In the spatial navigation task, participants had to imagine visiting their house each time they heard the spoken cue word "visit". More precisely, they had to list all objects inside a room of their house, and to switch to another room as soon as the previous one was fully listed. Participants had to imagine doing one of the imagery tasks, until they heard the word "relax". After hearing the word relax, participants had to stop their mental imagery for 30 seconds, until they heard the cue word again ("tennis" or "visit"). In each scan, five mental imagery periods were alternated with five rest periods. We defined the localizer when imagining playing tennis (motor localizer) as LocaM and the localizer when imagining home navigation (spatial navigation localizer) as LocaS.

2.2.2. Communication task

2.2.2.1. EEG

For the communication task of the EEG experiment, a question was asked after the second, forth, and sixth block of the command-following task. In this communication task, subjects attempted to answer three autobiographical yes-or-no questions (Q1, Q2, Q3) by modulating their brain activity. Questions remained similar from subject to subject ("do you have any brothers or sisters?", "Is your mother's maiden name…?", "Were you born in…?"). To answer these questions, participants were instructed to perform one of the two imagery tasks (hand imagery or toe imagery), one serving as "yes" and the other as "no". The imagery task serving as "yes" or "no" answer was balanced across subjects. The design was the same as for the command-following task, and subjects had to mentally respond to the question each time they heard a beep. A total of 15 consecutive tones, presented similarly to the EEG command-following task, were presented for each question.

2.2.2.2. fMRI

After the two runs of the fMRI command-following task, all participants underwent the fMRI communication-based task. In this communication task, subjects attempted to answer the same three questions as in the EEG communication task by modulating their brain activity. For half of the participants the answer "yes" was obtained by the motor imagery task, the other half having to perform the spatial navigation task. Investigators in charge of performing the fMRI task and data processing had not been informed of the answers before judging the fMRI data. The scanning procedure during the communication task was the same as the localizer, except for the beginning of each imagery period which was cued with the spoken word “answer”.

2.3. Data acquisition and analysis

2.3.1. EEG

2.3.1.1. EEG data acquisition

In the EEG experiment, all EEG channels were recorded using the OSG digital equipment (BrainRT; OSG bvba, Rumst, Belgium) with two Schwarzer AHNS epas 44 channels amplifiers (Natus, Munich, Germany). EEG signals were acquired from 64 electrodes at the positions of the 10/10 system using a 64 channel electrode cap (Easycap, EasycapGmbh, Ammersee, Germany). Sample frequency was set at 1000 Hz. Signal processing was performed using Cartool Software (http://brainmapping.unige.ch/Cartool.php). Epochs ranging from 1500ms prestimulus to 4500 ms poststimulus were extracted for each experimental condition and participant. Baseline was defined as the 500 ms period prior to stimulus onset. Individual data were then recalculated against the average reference and bandpass-filtered to 1-40 Hz. Data were visually inspected so as to reject epochs with blinks, eye movements or other sources of transient noise. Data at artifact electrodes from each participant were interpolated using a 3-dimensional spline algorithm (average: 6.25% interpolated electrodes). The average number of trials contributing to the healthy controls’ analyses was 55.2±3.5 for the right-hand condition and 55.2±3.2 for the toe condition. Similarly to the original study from (Cruse et al., 2011; Cruse et al., 2012a), only electrodes covering the motor area were selected. Here, 18 out of the 64 electrodes were chosen.

2.3.1.2. EEG data analysis

The classification was carried out using a support vector machine (SVM) designed using a linear kernel to which we fed a concatenation of the normalized average log-power values corresponding to each of the sliding windows of each channel. The supervised learning of the SVM was implemented for each subject using as training set the trials included in 6 out of the 8 blocks; the remaining 2 blocks were used for testing. The mean accuracy results obtained when constructing the training/testing sets were reported using two different methods:

- The testing set used for the SVM classification are the trials contained in a pair of consecutive blocks of different types; the training set consists of the trials included in the remaining blocks This is the technique proposed by (Cruse et al., 2011; Cruse et al., 2012a).
- The testing set used for the SVM classification are the trials contained in a pair of blocks of different types that, this time around, are not necessarily consecutive. This is the technique put forward by (Goldfine et al., 2013).

The ability of a given subject to adequately perform the requested imagery task was assessed by conducting a statistical test on the signal classification results whose H0 hypothesis is that the classification accuracy is 50%. When the testing sets consist of pairs of consecutive blocks of different types, the p-values for this test were computed as in (Cruse et al., 2011) where an independence hypothesis between trials (and hence between blocks) is invoked which allows us to declare that the number of correct answers follows under the null hypothesis a binomial distribution with parameters 0.5 and the number of trials. When the testing sets are constructed using necessarily consecutive pairs of blocks of different types, we computed the p-values by performing a permutation test, as proposed in (Goldfine et al., 2013), that consists of constructing an empirical distribution of accuracies obtained by considering the 34 different possible relabelings of the tasks attributed to the 8 blocks used for each subject. We also examined the consequences of applying a false-discovery rate (FDR) correction on the results, as proposed by (Goldfine et al., 2013). For the communication task, the signal acquisition and extraction was carried out using the same procedure as in the command-following task but the classification was implemented by using as training set all the blocks relative to the corresponding command-following task. The signal obtained during the answering of the questions was used as testing set.

2.3.2. fMRI

2.3.2.1. fMRI data acquisition

The functional MRI study was performed on a 3-Tesla (GE Healthcare Signa HDxt, Milwaukee, WI) MR system with a standard 40 mT/m gradient using blood–oxygen level-dependent (BOLD) fMRI. Foam cushions were used to minimize head movements within the coil. The experiment began with the acquisition of a high-resolution, T1-weighted, 3-dimensional anatomical scan (BRAVO sequence). This scan was acquired in 134 slices with 1mm x 1mm x 1mm resolution. Functional images were then obtained parallel to the anterior-posterior commissure line, covering the entire cerebrum (32 slices) using an echo planar imaging (EPI) sequence (slice thickness = 4 mm; In-plane resolution: 2.5mm x 2.5mm; TR = 2500 ms; TE = 35 ms; Flip Angle = 90°). The entire duration of the anatomical and functional scans lasted about 35min.

2.3.2.2. fMRI data analysis

Image time-series analysis was performed using BrainVoyager QX 2.1 (Brain Innovation, Maastricht, The Netherlands). The time-series were corrected for slice acquisition time, realigned with their corresponding T1 volumes, warped into standard space (Talairach and Tournoux, 1988), re-sampled into 3 mm isotropic voxels, motion-corrected using Levenberg-Marquarts's least square fit for six spatial parameters, highpass-filtered for removal of low frequency drifts, corrected voxel-wise for linear drifts, and spatially smoothed using a 8-mm full-width at half-maximum Gaussian kernel.

The general linear model (GLM) was computed from the z-normalized volume time courses. For mental imagery periods (in localizers as well as in communication tasks) and for rest periods, specific box-car time courses with a value of 1 corresponding to the period and values of 0 for the remaining time points were convolved with a theoretical hemodynamic response function (Boynton et al., 1996) and were entered as predictors into the design matrix of the study.

For single subject fixed-effect analyzes (first level analysis) a cluster size threshold yielding the equivalent of a multiple comparison correction significance level of $P < 0.05$ was used after voxel-wise thresholding at $P < 0.005$ ($t=2.86$) uncorrected. The BrainVoyager Cluster-Level Statistical Threshold Estimator plug-in estimated the overall significance level by determining the probability of false detection through Monte Carlo simulation (Forman et al., 1995) was used (with 1000 Monte Carlo iterations).

At group level (20 subjects) only localizer tasks were studied (LocaM and LocaS). Analyses were based on random effects (RFX) GLMs of the z-normalized volume time courses using a statistical threshold of $p<0.05$ corrected for multiple comparisons (Bonferroni corrected).

The regions of interest (ROIs) that were selected by (Boly et al., 2007; Monti et al., 2010), and expected to be activated in response to the localizers tasks are the left parahippocampal area (L-PPA) and the supplementary motor area (SMA). Therefore the first hypothesis was to find the L-PPA activated in the RFX analysis (at the group level) of the 20 spatial localizers and the SMA in the RFX analysis of the 20 motor localizers. We also checked that there was a good correspondence of the activations pattern with the one reported by Boly et al. To complete the group analysis expected to confirm this hypothesis, single-subject analyses aimed at detecting whether some subjects were unable to individually activate the expected ROI for one of or both the localizers. The analysis and classification of the activation patterns of the communication tasks allowing for the determination of subjects' responses were done according to the similarity measure. A similarity metric was computed to quantify how closely the activity in the regions of interest on each communication scan matched each localizer scan (see Supplementary Appendix of Monti et al. 2010 for details on how to use it). This method is based on the comparison, subject by subject, of t-values extracted from the L-PPA and the SMA in the communication scans Qi (i=1,...,3) and in both localizer scans.

3. Results

Table 1 shows a summary of the results of our 20 participants who performed both EEG and fMRI mental imagery tasks. At the bottom of the table the sensitivity of each imagery task with each neuroimaging method is presented. These results are discussed separately below.

--- INSERT TABLE 1 NEAR HERE ---

3.1. Command following

3.1.1. EEG

An example of synchronization/desynchronization in the mu band for each mental imagery task in three subjects is illustrated in figure (left). With the classification method described by (Cruse et al., 2011), we found a detection of command following in 60% of the healthy volunteers at the alpha=.05 level, and 70% at the alpha=.1 level. After the statistical correction proposed by (Goldfine et al., 2013), command following was observed in only 6 out of the 20 participants without an FDR correction, leading to a sensitivity of 30%. After FDR correction, a significant command following was only found in three subjects (S1, S11, S13), which means that the sensitivity of the EEG protocol drops drastically to 15%.

--- INSERT FIGURE NEAR HERE ---

3.1.2. fMRI

The group analysis revealed similar activation patterns as those previously found in the literature. The regions of interest were activated during imagery: the SMA/pre-SMA complex is strongly activated during the LocaM (t=10.91, p<.05 Bonferroni corrected) and the L-PPA is strongly activated during the LocaS (t=11.60, p<.05

Bonferroni corrected). To investigate whether the differences observed at group level reflect the same patterns of cerebral activations in all healthy participants, results from first-level analysis were studied. Figure (right) shows an example of patterns of activations in three subjects.

When considering the SMA and L-PPA alone as markers of LocaM and LocaS, these exact patterns of activations were not observed in 15% of participants. Indeed, 1 out of 20 subjects did not activate the L-PPA during spatial imagery (LocaS), and two other subjects did not activate the SMA during motor imagery (LocaM). Interestingly, these areas of interest were not activated at lower statistical thresholds either. The sensitivity of the method was calculated by considering subjects who activated both the SMA during the motor imagery task and the L-PPA during the spatial imagery task as true positives. In our results, the sensitivity of the fMRI protocol was 85%. The L-PPA was also activated in 25% of participants (5/20) during motor imagery, and 90% (18/20) of participants activated the SMA during spatial imagery.

3.1.3. Combination of EEG and fMRI

Eighteen of our subjects (90%) showed distinct command following with at least one of the two methods (table 2). After the statistical correction proposed by (Goldfine et al., 2013), a significant command following was found only with EEG in one subject, and only with fMRI in 12 subjects. No significant command-following task was found either in EEG or in fMRI for two subjects.

--- INSERT TABLE 2 NEAR HERE ---

3.2. Communication

The sensitivity of both EEG and fMRI communication tasks was calculated by counting as true positives cases in which all (3/3) answers were correctly detected in a given participant.

3.2.1. EEG

With EEG, only 62 percent of answers were correctly detected, which was not significantly different from chance level (chi^2 = 1.656, n.s.). However, for the six subjects who showed good results in the command following task, 16 answers out of the 18 questions were correctly interpreted. The sensitivity of the EEG communication task was of 30%.

3.2.2. fMRI

The results obtained with the similarity metric proposed by (Monti et al., 2010), which allowed for a 100% decoding in 2010, gave an error rate of 20% in our study. The number of correct answers was significantly above chance level (chi^2 = 19.780, $p<.0001$). Of the 17 participants who activated specifically the SMA during motor imagery and the L-PPA during spatial imagery (in the command following task), a 100% correct detection was not obtained. The sensitivity of fMRI communication task was of 12/(12+8) = 60%.

3.2.3. Combination of EEG and fMRI

A 100% correct detection rate in both EEG and fMRI communication tasks (table 2) was only found in four participants. No correct detection of answers could be found either in EEG or fMRI in 6 of our participants.

3.3. Questionnaires

The analyses of the questionnaire given at the end of both experiments revealed that 10 subjects found that mental imagery tasks were easier to perform with fMRI, whereas only 3 subjects had a preference for EEG. Concerning the environment in which each exam was performed, 10 subjects expressed a preference for EEG, 2 subjects for the fMRI environment, and 8 subjects were not disturbed by either condition. Overall, 9 subjects preferred the fMRI exam over EEG, 7 subjects preferred EEG, and 4 subjects found both exams as equivalent.

4. Discussion

EEG and fMRI are two widely used techniques to assess awareness in disorders of consciousness such as coma, UWS, or MCS. Because of the different nature of the signal recorded and the constraints inherent to each method, their respective role is usually clearly attributed to specific categories of patients. Here, the similar mental imagery protocols used in EEG and fMRI raise questions about the relative place of each method. The results presented here reveal that more than a substitute, each method should be viewed as complementary.

4.1. Command-following task

In the EEG command-following task designed by (Cruse et al., 2011), the presence of awareness was detected in 60% of our healthy volunteers at the alpha=.05 level, which is close to the results obtained in the control group of the original study. However, a very low sensitivity was obtained when the analyses were corrected with the more conservative statistical models suggested by (Goldfine et al., 2013),.

Significant detection of willful command following was then found in only 6 (without FDR correction) or 3 (after FDR correction) out of the 20 subjects. These high discrepancies of sensitivity which depend on the statistical method used are a current concern in the interpretation of EEG data. Additionally, the test used for statistical analysis by (Cruse et al., 2011) involves assumptions that cannot be justified and that, as it is shown in (Henriques et al., 2014), are likely to be violated which hence may produce misleading results. Indeed, in this protocol, the cross-validation design is a particularly delicate issue that has already been pointed out (Goldfine et al., 2013; Noirhomme et al., 2014) and that shows the strong influence of the temporal dependence between the test-set blocks on the classification accuracy. In recent studies avoiding this limitation, promising results have been found in healthy volunteers but also in some patients with DOCs (Gibson et al., 2013; Holler et al., 2013) but their replication on a larger scale is needed before clearly establish whether they are sensitive enough to reliably capture awareness. By contrast, the conservatism of FDR correction is more a matter of debate. It has been argued that this conservatism could be gauged by the importance of the results (Cruse et al., 2014). Whatever the correction chosen, the clinical interpretation of such results should be taken with extreme care. With the use of neuroimaging methods, the determination of a specific cerebral response relies on statistical analyses only. The relatively low detection levels attained with the available EEG technology overplay the role of the significance levels chosen for the different statistical tests that are used to assess the awareness condition of the patient. In these circumstances, the choice of a high or low significance level may have dramatic consequences that can potentially harm patients and families. Therefore, even though the collective final goal of the research field in which our work takes place is the development of factual and quantifiable diagnosis tools, the current state of the art in the EEG detection technology imposes extreme caution at the time of making decisions based on the associated classification results.

To our surprise, the replication of the fMRI command following task from (Owen et al., 2006) did not lead to a 100% sensitivity like in the original studies when performed in healthy volunteers (Boly et al., 2007; Monti et al., 2010; Owen et al., 2006) or in patients with locked-in-syndrome (Stender et al., 2014) as control subjects. A small proportion of our healthy subjects failed to show a clear brain response in at least one of the two fMRI mental imagery tasks, which confirms recent results obtained with the same paradigm (Fernandez-Espejo et al., 2014).Here, we found robust activity in the supplementary motor area in 18 out of 20 subjects when they were instructed to imagine playing tennis. When the instruction was to imagine moving around their houses, 19 out of the 20 subjects activated the parahippocampal cortices. For the command following task, if we consider as aware the participants who activated specifically both areas of interest (Monti et al., 2010; Owen et al., 2006), 17 out of the 20 subjects would have been correctly classified. Although we did not reach a sensitivity of 100%, sensitivity remains high at 85% which is far above the one obtained with the EEG command-following task.

A major result consisted in the detection of awareness with only one of the neuroimagery methods in some subjects. Several subjects who could not follow command with EEG did perform well with fMRI, and the contrary was also observed for one participant (subject 6). This subject did not activate the SMA during the motor imagery task in fMRI, but his mental imagery tasks were accurately classified in EEG. Thus, to improve the detection of volitional awareness, both neuroimaging protocols should be used together. Currently, fMRI is mostly used to corroborate the presence of awareness detected in EEG in patients with DOCs (Cruse et al., 2013) (Cruse et al., 2013) or to give additional information to the EEG results (Chennu et al., 2013). Only one recent study has compared the activations obtained with both EEG and fMRI mental imagery tasks on a group of patients with DOCs (Gibson et al., 2014). Large divergences of responses were found between both methods in these patients, confirming the importance of employing multiple modalities. As we show here, these differences may not necessarily originate from the pathology, such as varying levels of awareness between both exams, but from the design of the experiment and thus should be taken with extreme caution. The fact that two of our healthy subjects did not successfully respond to EEG and fMRI command-following tasks underlines that even in aware subjects, signs of awareness might remain undetected with current neuroimaging protocols.

4.2. Communication task

Similarly to what has been done with fMRI, one of our objectives was to explore whether the EEG command-following tasks could efficiently be used as a communication tool. We obtained unsatisfactory results that were certainly impacted by the low sensitivity of the command-following task. At the group level, the performance in detecting answers did not differ from chance. However, for the 6 subjects who showed significant activations in the command following task (without FDR correction), the results were nevertheless satisfactory with the correct detection of responses for 16 out of 18 questions. The possibility of using EEG mental imagery tasks as a communication tool might thus be possible but need to be further explored, especially by adding a larger number of questions for a given subject to ensure the reliability of the task. It is also possible to use other approaches to communicate with EEG, for example through the use of evoked potentials, or the regulation of slow cortical potentials, or others (Falkenstein et al., 1994; Furdea et al., 2009; Ross et al., 2004). First promising results have been reported with the use of auditory attention to communicate with patients with DOCs and will have to be thoroughly investigated (Lule et al., 2013).

In fMRI, it has already been shown that some brain injured patients who proved able to communicate verbally or through head movements may be unable to perform the fMRI communication task efficiently (Bardin et al., 2011). Here, we show that even healthy controls may fail to perform the task since sensitivity only reached a value of 60% and that correct detection was only obtained in 12 subjects. Only 80% of all the correct answers could be detected with this method. Furthermore, subjects who did perform well at the command following task did not get better results than those who failed, contrary to what was found with EEG. The reasons why we did not detect 100% of correct responses as in the control group of the original study of Monti et al. (2010) are difficult to determine. For each question, the percentage of similarity between the correct answers and the corresponding command following task was of an average of 65% , which is below the average similarity of 82% found in the group of 16 healthy volunteers tested in the study of (Monti et al., 2010). Of the reasons that might have impacted on the results, one must consider that slight differences were present between our analyses and those of (Monti et al., 2010). For example, at the individual level we estimated the overall significance level by determining the probability of false detection through the Monte Carlo simulation. It incorporates the observation that neighboring voxels often activate in clusters, as hypothesized with Gaussian random field (RF). However the chosen approach does not require substantial spatial smoothing contrarily to RF, which is known not to perform well in some settings when theoretical approximations are not accurate (Hayasaka and Nichols, 2003; Smith and Nichols, 2009). The reduced sensitivity observed in our study may also emerge from the heterogeneity of the chosen population in terms of age, sex and education level, or by slightly different instructions from those given in the original study. Whatever the reason underlying this reduction of performances, our results show that the generalization of the fMRI communication method to populations of patients with DOCs asks for a good understanding of the methodology and an extreme care in the interpretation of results in practice. Interestingly, 70% of the subjects (14/20) responded correctly to 100% of the questions with at least one of the two methods, which increases the possibility to communicate with them. Once again, it is rather the combination of both neuroimaging methods than a mere substitution of one by the other that should be considered in future investigations.

4.3. Relative place of EEG and fMRI

Our results indicate that fMRI and EEG protocols from (Cruse et al., 2011) and (Monti et al., 2010) may not be sensitive enough to eliminate the risk of not capturing awareness in aware subjects. This is all the more important that our EEG and fMRI data obtained on healthy volunteers were not distorted by ocular or muscular artifacts, which is far from being the case with patients. Furthermore, is it not uncommon to reject data from more than 40% of patients performing mental imagery tasks with fMRI, because of spontaneous movements during the recording (Stender et al., 2014). As a consequence, a special care should be taken in case of negative findings obtained on patients with DOCs with one of the two methods. This could not only reflect a deficiency in one of the cognitive abilities of the patients, but also be attributed to variations in arousal, or reflect a lack of sensitivity of the neuroimaging method used.

Since fMRI is a technique materially difficult to generalize in terms of cost, availability, or patients movements (Stender et al., 2014), it is tempting to consider EEG as a substitute for fMRI or as a first screening procedure. However, since the EEG command-following task showed a far reduced sensitivity compared to fMRI, its use in routine care cannot be fully recommended. The present results show that combining both EEG and fMRI mental imagery tasks improves the detection of awareness and the possibility to establish a way to communicate. If both EEG and fMRI methods are available in a coma center, the conjoint use of these methods on a same patient would increase the probability of detecting signs of awareness.

In healthy volunteers, carrying out separate mental imagery sessions with EEG and fMRI is possible, but a potential drawback to use them in patients with DOCs is their fluctuating degree of awareness, which renders the assessment of awareness even more difficult. A solution to this issue may be to perform EEG and fMRI recordings simultaneously. With combined EEG-fMRI, the investigation of the neural activity at a high temporal and spatial resolution could be achieved. Since several studies have shown the presence of an (inverse) correlation between alpha power and the BOLD signal (Goldman et al., 2002; Mo et al., 2013), the variations of signal induced by mental imagery tasks could thus be simultaneously observed with each neuroimaging method.

Of importance is also to continue developing new protocols and new statistical methods specific to each neuroimaging method because one or more methods may not be appropriate to a given patient with DOC. For example, in the sample of 14 patients recruited by (Gibson et al., 2014), some of them were ineligible for the fMRI and/or the EEG evaluation and only 6 could complete both procedures. Promising results have recently been found in EEG as well as in fMRI (Cruse et al., 2012b; Gibson et al., 2013; Goldfine et al., 2011; Holler et al., 2013; Naci and Owen, 2013; Schnakers et al., 2009). Better sensitivity may be observed with these protocols than with the one used in the present study, but since these protocols were performed by different teams in different groups of subjects it is difficult to determine precisely.

5. Conclusion

In conclusion, the results presented here show that mental imagery tasks alone are not sensitive enough to detect awareness in fully aware subjects. This confirms that negative findings in these neuroimaging protocols should never be considered as evidence of lack of awareness in patients with DOCs, but rather that the methods may not be sensitive enough to capture awareness. Although the detection of awareness was not possible for some subjects with both neuroimaging methods, our results give strong evidence that they should be used in complement to reduce the presence of negative findings. It is also important to keep in mind that current approaches require a good understanding of the methodology and of the statistical models to be used to interpret the collected data. Therefore, standardized behavioral assessments are still the standard in the evaluation of consciousness in patients with DOCs.

Contributors:

The following authors conceived and designed the experiments: DG, AC, EC, EH, TM, RA, LP. Performed the experiments: DG, AC. Analyzed the data: DG JH AC LG JPO, LP. Contributed to the writing of the manuscript: DG, JH, AC, LG, JPO, GB, RA, LP.

Acknowledgments:

This work was supported by the 2011 Translational clinical research program from the French Ministry of Health and Inserm.

Bardin, J.C., Fins, J.J., Katz, D.I., Hersh, J., Heier, L.A., Tabelow, K., Dyke, J.P., Ballon, D.J., Schiff, N.D., and Voss, H.U. (2011). Dissociations between behavioural and functional magnetic resonance imaging-based evaluations of cognitive function after brain injury. Brain *134*, 769-782.

Bekinschtein, T.A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., and Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. Proc Natl Acad Sci U S A *106*, 1672-1677.

Boly, M., Coleman, M.R., Davis, M.H., Hampshire, A., Bor, D., Moonen, G., Maquet, P.A., Pickard, J.D., Laureys, S., and Owen, A.M. (2007). When thoughts become action: an fMRI paradigm to study volitional brain activity in non-communicative brain injured patients. Neuroimage *36*, 979-992.

Boynton, G.M., Engel, S.A., Glover, G.H., and Heeger, D.J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci *16*, 4207-4221.

Chennu, S., Finoia, P., Kamau, E., Monti, M.M., Allanson, J., Pickard, J.D., Owen, A.M., and Bekinschtein, T.A. (2013). Dissociable endogenous and exogenous attention in disorders of consciousness. Neuroimage Clin *3*, 450-461.

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T.A., Fernandez-Espejo, D., Pickard, J.D., Laureys, S., and Owen, A.M. (2011). Bedside detection of awareness in the vegetative state: a cohort study. Lancet *378*, 2088-2094.

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T.A., Fernandez-Espejo, D., Pickard, J.D., Laureys, S., and Owen, A.M. (2013). Reanalysis of "Bedside detection of awareness in the vegetative state: a cohort study" - Authors' reply. Lancet *381*, 291-292.

Cruse, D., Chennu, S., Chatelle, C., Fernandez-Espejo, D., Bekinschtein, T.A., Pickard, J.D., Laureys, S., and Owen, A.M. (2012a). Relationship between etiology and covert cognition in the minimally conscious state. Neurology *78*, 816-822.

Cruse, D., Chennu, S., Fernandez-Espejo, D., Payne, W.L., Young, G.B., and Owen, A.M. (2012b). Detecting awareness in the vegetative state: electroencephalographic evidence for attempted movements to command. PLoS One *7*, e49933.

Cruse, D., Gantner, I., Soddu, A., and Owen, A.M. (2014). Lies, damned lies and diagnoses: Estimating the clinical utility of assessments of covert awareness in the vegetative state. Brain Inj *28*, 1197-1201.

Falkenstein, M., Hohnsbein, J., and Hoormann, J. (1994). Effects of choice complexity on different subcomponents of the late positive complex of the event-related potential. Electroencephalogr Clin Neurophysiol *92*, 148-160.

Fernandez-Espejo, D., Norton, L., and Owen, A.M. (2014). The clinical utility of fMRI for identifying covert awareness in the vegetative state: a comparison of sensitivity between 3T and 1.5T. PLoS One *9*, e95082.

Fernandez-Espejo, D., and Owen, A.M. (2013). Detecting awareness after severe brain injury. Nat Rev Neurosci *14*, 801-809.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., and Noll, D.C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med *33*, 636-647.

Furdea, A., Halder, S., Krusienski, D.J., Bross, D., Nijboer, F., Birbaumer, N., and Kubler, A. (2009). An auditory oddball (P300) spelling system for brain-computer interfaces. Psychophysiology *46*, 617-625.

Ghaem, O., Mellet, E., Crivello, F., Tzourio, N., Mazoyer, B., Berthoz, A., and Denis, M. (1997). Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. Neuroreport *8*, 739-744.

Gibson, R.M., Chennu, S., Owen, A.M., and Cruse, D. (2013). Complexity and familiarity enhance single-trial detectability of imagined movements with electroencephalography. Clin Neurophysiol.

Gibson, R.M., Fernandez-Espejo, D., Gonzalez-Lara, L.E., Kwan, B.Y., Lee, D.H., Owen, A.M., and Cruse, D. (2014). Multiple tasks and neuroimaging modalities increase the likelihood of detecting covert awareness in patients with disorders of consciousness. Front Hum Neurosci *8*, 950.

Goldfine, A.M., Bardin, J.C., Noirhomme, Q., Fins, J.J., Schiff, N.D., and Victor, J.D. (2013). Reanalysis of "Bedside detection of awareness in the vegetative state: a cohort study". Lancet *381*, 289-291.

Goldfine, A.M., Victor, J.D., Conte, M.M., Bardin, J.C., and Schiff, N.D. (2011). Determination of awareness in patients with severe brain injury using EEG power spectral analysis. Clin Neurophysiol *122*, 2157-2168.
Goldman, R.I., Stern, J.M., Engel, J., Jr., and Cohen, M.S. (2002). Simultaneous EEG and fMRI of the alpha rhythm. Neuroreport *13*, 2487-2492.
Hayasaka, S., and Nichols, T.E. (2003). Validating cluster size inference: random field and permutation methods. Neuroimage *20*, 2343-2356.
Henriques, J., Gabriel, D., Grigoryeva, L., Haffen, E., Moulin, T., Aubry, R., Pazart, L., and Ortega, J.P. (2014). Protocol Design Challenges in the Detection of Awareness in Aware Subjects Using EEG Signals. Clin EEG Neurosci.
Holler, Y., Bergmann, J., Thomschewski, A., Kronbichler, M., Holler, P., Crone, J.S., Schmid, E.V., Butz, K., Nardone, R., and Trinka, E. (2013). Comparison of EEG-features and classification methods for motor imagery in patients with disorders of consciousness. PLoS One *8*, e80479.
Lotze, M., Montoya, P., Erb, M., Hulsmann, E., Flor, H., Klose, U., Birbaumer, N., and Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. J Cogn Neurosci *11*, 491-501.
Lule, D., Noirhomme, Q., Kleih, S.C., Chatelle, C., Halder, S., Demertzi, A., Bruno, M.A., Gosseries, O., Vanhaudenhuyse, A., Schnakers, C., *et al.* (2013). Probing command following in patients with disorders of consciousness using a brain-computer interface. Clin Neurophysiol *124*, 101-106.
Mo, J., Liu, Y., Huang, H., and Ding, M. (2013). Coupling between visual alpha oscillations and default mode activity. Neuroimage *68*, 112-118.
Monti, M.M., Vanhaudenhuyse, A., Coleman, M.R., Boly, M., Pickard, J.D., Tshibanda, L., Owen, A.M., and Laureys, S. (2010). Willful modulation of brain activity in disorders of consciousness. N Engl J Med *362*, 579-589.
Naci, L., and Owen, A.M. (2013). Making every word count for nonresponsive patients. JAMA Neurol *70*, 1235-1241.
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*, 687-694.
Owen, A.M., Coleman, M.R., Boly, M., Davis, M.H., Laureys, S., and Pickard, J.D. (2006). Detecting awareness in the vegetative state. Science *313*, 1402.
Pfurtscheller, G., Brunner, C., Schlogl, A., and Lopes da Silva, F.H. (2006). Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. Neuroimage *31*, 153-159.
Pfurtscheller, G., and Lopes da Silva, F.H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol *110*, 1842-1857.
Ross, B., Picton, T.W., Herdman, A.T., and Pantev, C. (2004). The effect of attention on the auditory steady-state response. Neurol Clin Neurophysiol *2004*, 22.
Schnakers, C., Perrin, F., Schabus, M., Hustinx, R., Majerus, S., Moonen, G., Boly, M., Vanhaudenhuyse, A., Bruno, M.A., and Laureys, S. (2009). Detecting consciousness in a total locked-in syndrome: an active event-related paradigm. Neurocase *15*, 271-277.
Schnakers, C., Perrin, F., Schabus, M., Majerus, S., Ledoux, D., Damas, P., Boly, M., Vanhaudenhuyse, A., Bruno, M.A., Moonen, G., *et al.* (2008). Voluntary brain processing in disorders of consciousness. Neurology *71*, 1614-1620.
Smith, S.M., and Nichols, T.E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage *44*, 83-98.
Stender, J., Gosseries, O., Bruno, M.A., Charland-Verville, V., Vanhaudenhuyse, A., Demertzi, A., Chatelle, C., Thonnard, M., Thibaut, A., Heine, L., *et al.* (2014). Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study. Lancet *384*, 514-522.
Talairach, J., and Tournoux, P. (1988). Co-Planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System : An Approach to Cerebral Imaging (New York, Thieme Medical Publishers).

Table 1: Comparison of EEG and fMRI classification accuracies for each subject. EEG accuracies from Goldfine are displayed without FDR correction. In the fMRI command following task, sensitivity is calculated by considering as true positives subjects for whom a significant detection was obtained during both the motor and the spatial imagery. In the EEG and fMRI communication task, true positives are represented by subjects for whom 3/3 correct answers were detected. The sensitivity of EEG and fMRI tasks is noted at the bottom of the table.

			command-following task						communication task							
			EEG				fMRI (Monti et al.,2010)		EEG				fMRI (Monti et al., 2010)			
			Cruse et al. (2011; 2012)		Goldfine et al. (2013)		p-value<0.005		answer accuracy (%)			Total correct	answer accuracy (%)			Total correct
Subject No.	Age	Sex	Accuracy (%)	p-value	Accuracy (%)	p-value	motor	spatial	Q1	Q2	Q3	answers	Q1	Q2	Q3	answers
1	25	F	**71**	**0.000**	60	0.118	**yes**	**yes**	**73**	50	**75**	2 /3	**70**	49	43	1 /3
2	29	M	**64**	**0.001**	60	0.088	**yes**	**yes**	**100**	50	36	1 /3	**64**	**68**	**66**	**3 /3**
3	35	F	56	0.116	49	0.500	**yes**	**yes**	**67**	27	47	1 /3	**66**	**67**	**69**	**3 /3**
4	29	F	**78**	**0.000**	**72**	**0.029**	**yes**	**yes**	**67**	**79**	**60**	**3 /3**	**55**	**78**	**73**	**3 /3**
5	35	M	**65**	**0.001**	57	0.206	**yes**	**yes**	**93**	13	47	1 /3	**66**	**77**	**65**	**3 /3**
6	28	M	**61**	**0.009**	**62**	**0.000**	no	**yes**	**80**	**71**	23	2 /3	**69**	**76**	**65**	**3 /3**
7	32	M	54	0.226	41	0.677	**yes**	**yes**	**73**	36	39	1 /3	27	32	**56**	1 /3
8	27	F	**71**	**0.000**	65	0.088	**yes**	**yes**	40	31	33	0 /3	**53**	**95**	**64**	**3 /3**
9	32	F	**65**	**0.001**	60	0.059	**yes**	**yes**	**57**	27	50	1 /3	**68**	44	**63**	2 /3
10	26	M	43	0.945	41	0.735	**yes**	no	7	20	**92**	1 /3	**85**	**57**	**52**	**3 /3**
11	27	M	**75**	**0.000**	**76**	**0.000**	**yes**	**yes**	**54**	**87**	**79**	**3 /3**	**95**	**89**	**68**	**3 /3**
12	34	M	58	0.051	53	0.294	**yes**	**yes**	**58**	29	**75**	2 /3	**82**	**88**	**72**	**3 /3**
13	66	F	**73**	**0.000**	**68**	**0.000**	**yes**	**yes**	**93**	**71**	**73**	**3 /3**	46	**53**	**72**	2 /3
14	29	F	50	0.576	44	0.735	**yes**	**yes**	**64**	**79**	**80**	**3 /3**	**55**	**74**	**75**	**3 /3**
15	47	F	45	0.879	30	1.000	**yes**	**yes**	23	23	**67**	1 /3	45	**86**	**92**	2 /3
16	28	F	58	0.077	54	0.265	**yes**	**yes**	**55**	43	**55**	2 /3	**90**	**80**	**66**	**3 /3**
17	55	F	**66**	**0.001**	59	0.177	no	**yes**	**71**	**100**	**92**	**3 /3**	**82**	17	38	1 /3
18	42	M	48	0.683	38	0.824	**yes**	**yes**	40	**64**	**79**	2 /3	**68**	**85**	34	2 /3
19	35	F	**67**	**0.000**	**64**	**0.029**	**yes**	**yes**	**82**	**53**	**92**	**3 /3**	**63**	**63**	**64**	**3 /3**
20	51	M	**60**	**0.021**	**62**	**0.029**	**yes**	**yes**	27	**69**	**80**	2 /3	48	21	**56**	1 /3
Mean	35.6		61%		56%				61%	51%	64%	62%	65%	65%	63%	80%
Sensitivity			**60%**		**30%**		**85%**					**30%**				**60%**

Table 2: Proportion of subjects responding to one, both or none of the imaging methods in the command-following and in the communication task. In the EEG command-following task, main values are obtained from Goldfine without FDR correction. In brackets are the subjects obtained from Cruse.

Command-following task (n=20)

		fMRI success	fMRI failure
EEG	success	5 (10)	1 (2)
EEG	failure	12 (7)	2 (1)

Communication task (n=20)

		fMRI success	fMRI failure
EEG	success	4	2
EEG	failure	8	6

Figure: EEG and fMRI patterns of activations obtained during the command-following task in three participants. Left: Topographic representation of the cerebral synchronization/desynchronization observed in the mu band while subjects had to imagine squeezing their right-hand (hand) or wiggling the toes of their feet (feet) in the EEG task. For this illustration, the calculation of synchronization percentages follow the study from (Pfurtscheller and Lopes da Silva, 1999). The area within the red border corresponds to the motor area, which is the only one considered for classification. Please note a desynchronization of the left motor area in S11 and S6 while the subjects were imagining squeezing their right hand. Right: Patterns of cerebral activations in the fMRI task during motor imagery (motor) and spatial navigation (spatial). In all subjects, an activation of the parahippocampal area is observed during spatial navigation. In S06, no activation of the SMA is observed during motor imagery. Please note that in order to enable a comparison with EEG, the left side of the brain is shown on the left side of each MRI slide.