# Neural responses to heartbeats detect residual signs of consciousness during resting state in post-comatose patients

Abbreviated title: Responses to heartbeats detect residual consciousness

Diego Candia-Rivera[1*], Jitka Annen[2, 3*], Olivia Gosseries[2, 3], Charlotte Martial[2, 3], Aurore Thibaut[2, 3], Steven Laureys[2, 3#], Catherine Tallon-Baudry[1#]

[1] Laboratoire de Neurosciences Cognitives et Computationnelles, Département d'Etudes Cognitives, École Normale Supérieure, INSERM, Université PSL, Paris, France.

[2] GIGA-Consciousness, Coma Science Group, University of Liège, Liège, Belgium.

[3] Centre du Cerveau, University Hospital of Liège, Liège, Belgium.

* co-first authors

# co-last authors

Submitting and corresponding author: Diego Candia-Rivera (diego.candia.r@ug.uchile.cl)

Pages: 47 Figures: 4 Tables: 4

Words, abstract: 247 Words, Introduction: 683 Words, discussion: 1595

Conflict of interest: The authors declare no competing financial interests

**Acknowledgements**

The authors thank the patients and their legal guardians for their consent to perform this study and the whole staff from the Neurology, Radiodiagnostic and Nuclear Medicine departments, University Hospital of Liege. We are highly grateful to the members of the Liège Coma Science Group for their assistance in clinical evaluations.

C.T.B. is supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 670325, Advanced grant BRAVIUS), by a senior fellowship from the Canadian Institute For

Advance Research (CIFAR) program in Brain, Mind and Consciousness, as well as by ANR-17-EURE-0017. AT, OG and SL are researchers at FRS-FNRS. The study was further supported by the University and University Hospital of Liege, the Belgian National Funds for Scientific Research (FRS-FNRS), the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3), the European Space Agency (ESA) and the Belgian Federal Science Policy Office (BELSPO) in the framework of the PRODEX Programme, "Fondazione Europea di Ricerca Biomedica", the Bial Foundation, the Mind Science Foundation, the fund Generet, the King Baudouin Foundation, and DOCMA project [EU-H2020-MSCA–RISE–778234].

## Abstract

The neural monitoring of visceral inputs might play a role in first-person perspective, i.e. the unified viewpoint of subjective experience. In healthy participants, how the brain responds to heartbeats, measured as the heartbeat-evoked response (HER), correlates with perceptual, bodily, and self-consciousness. Here we show that HERs in resting-state EEG data distinguishes between post-comatose male and female human patients (n=68, split into training and validation samples) suffering from the unresponsive wakefulness syndrome and patients in minimally conscious state with high accuracy (random forest classifier, 87% accuracy, 96% sensitivity and 50% specificity in the validation sample). Random EEG segments not locked to heartbeats were useful to predict (un)consciousness, but HERs were more accurate, indicating that HERs provide specific information on consciousness. HERs also led to more accurate classification than heart rate variability. HER-based consciousness scores correlate with glucose metabolism in the default mode network node located in the right superior temporal sulcus, as well as with the right ventral occipito-temporal cortex. These results were obtained when consciousness was inferred from brain glucose metabolism measured with Positron Emission Topography. HERs reflected the consciousness diagnosis based on brain metabolism better than the consciousness diagnosis based on behavior (Coma Recovery Scale-Revised, 77% validation accuracy). HERs thus seem to capture a capacity for consciousness that does not necessarily translate into intentional overt behavior. These results confirm the role of HERs in consciousness, offer new leads for future bedside testing, and highlight the importance of defining consciousness and its neural mechanisms independently from behavior.

## Significance statement

Detecting consciousness without relying neither on overt behavior nor on asking to mentally perform a specific task is both a fundamental issue pertaining to the nature of consciousness, and a clinical challenge. Here we show that the transient brain response elicited at each heartbeat captures residual consciousness in the resting-state EEG of post-comatose patients. The results show that brain responses to an internal bodily signal might help specify the gray zone of consciousness, i.e., the fleeting conscious feelings that are not necessarily associated with the performance of a specific task nor translate into behavioral outputs.

## Introduction

We recently proposed that the neural monitoring of signals ascending from the heart and gastro-intestinal tract plays an important role in consciousness (Park and Tallon-Baudry, 2014; Tallon-Baudry et al., 2018; Azzalini et al., 2019). Visceral inputs are intrinsically private and might thus be self-specifying. More precisely, the neural responses to visceral inputs would contribute to conscious experience through first-person perspective, or the bodily-centered viewpoint from which we subjectively experience both the environment and inner mental life (Blanke and Metzinger, 2009). This represents a core component of the simplest, but also most elusive, aspect of consciousness. A number of experimental results in healthy adult participants support this hypothesis. How the brain transiently responds to heartbeats can be experimentally measured by averaging EEG or MEG data time-locked to heartbeats, to generate the heartbeat-evoked response (HER) (Schandry et al., 1986). In healthy participants, HERs predict perceptual consciousness (Park et al., 2014; Al et al., 2020), reflect bodily consciousness (Park et al., 2016; Sel et al., 2017) and the self vs. other distinction (Babo-Rebelo et al., 2019). HERs also co-vary with the self-relatedness of spontaneous thoughts, as rated by participants themselves (Babo-Rebelo et al., 2016a, 2016b), suggesting that HERs do capture a component of consciousness not related to task performance. We thus hypothesized that heartbeat-evoked responses and their fluctuations could be a marker of consciousness even in the absence of overt behavior or mental response to instructions in the resting-state EEG of post-comatose patients with disorders of consciousness.

Probing consciousness in the absence of overt behavior is theoretically motivated but remains an experimental and clinical challenge. Consciousness is defined by the existence of subjective experience and inner mental life (Chalmers, 1995; Block, 2005), that does not necessarily translate into overt behavior (Tsuchiya et al., 2015) – in other words, subjective experience is necessary for consciousness while overt behavior is not. If follows that

experimental work should focus on the neural mechanisms giving rise to conscious experience, rather than on the cognitive processes required for report in healthy participants (Frässle et al., 2014). In clinical practice, the threshold for (un)consciousness is currently placed between patients showing eye opening but only reflex-like responses to the environment (Unresponsive Wakefulness Syndrome; UWS (Laureys et al., 2010)) and patients with fluctuating but reproducible signs of non-reflex behavior (Minimally Conscious State; MCS (Giacino et al., 2002)). The clinical rationale is thus based solely on behavioral signs of consciousness (Bayne et al., 2017). In addition to the criticisms raised above, measuring consciousness from behavior is an issue in patients who might suffer from motor, sensory or cognitive deficits, or in patients lacking the motivation or attentional resources to respond to the command. Consciousness indices based on brain activity and which do not require patients' active participation have been developed and validated, such as neural responses to magnetic stimulation applied transcranially (Casali et al., 2013; Casarotto et al., 2016), or cerebral glucose metabolism at rest obtained with $^{18}$F-fluorodeoxyglucose positron emission tomography (FDG-PET) (Tommasino et al., 1995; Rudolf et al., 1999; Laureys et al., 2004; Nakayama et al., 2006; Thibaut et al., 2012; Gosseries et al., 2014; Stender et al., 2014, 2015). While behavioral assessments remain the clinical standard (Giacino et al., 2018; Kondziella et al., 2020), brain imaging provides complementary information in patients without command following, and is recommended by the European Academy of Neurology (Kondziella et al., 2020). In particular, FDG-PET seems useful to detect residual consciousness, or predict potential for consciousness recovery, from brain metabolism in patients without any behavioral sign of consciousness (Stender et al., 2014).

We tested whether HERs could reliably detect residual consciousness in post-comatose patients with disorders of consciousness using resting-state high-density EEG data (n=68; training sample n=38; validation sample n=30, see Table 1). To probe consciousness

independently from behavior, we used the neuroimaging diagnosis of consciousness based on resting-state brain glucose uptake obtained with FDG-PET. EEG was measured during the FDG uptake phase. We further explored cases where the PET-based diagnosis was not congruent with the behavioral diagnosis, obtained from repeated standardized clinical assessments using the Coma Recovery Scale-Revised (CRS-R, (Giacino et al., 2004; Wannez et al., 2017)).

# Material and Methods

## Patients

Patients included in this study were referred to the University Hospital of Liège between January 2008 and October 2015 for a one-week assessment of their state of consciousness. This included the acquisition of resting-state FDG-PET, high-density EEG data and CRS-R assessments.

Included patients presented a prolonged (at least 28 days) disorder of consciousness after severe brain damage, based on international guidelines and repeated CRS-R assessments (Giacino et al., 2018). Exclusion criteria were being under-age or suffering from pre-existing psychological or neurological diseases, sedative drugs, FDG-PET and EEG contraindications. For the purpose of the present study we trained classifiers on the consciousness diagnosis based on the assessment of FDG-PET data, hereafter termed PET-based diagnosis as in (Stender et al., 2014). More details on the PET-based diagnosis are provided in the section "PET data and correlation with consciousness scores". Similar analysis were performed using the clinical diagnosis based on the best assessment with the Coma Recovery Scale-Revised (CRS-R) out of at least five assessments (Wannez et al., 2017), hereafter termed CRS-R diagnosis. The CRS-R assessment is based on at least 5 assessments performed on separate days within the one-week hospitalization, including one CRS assessment systematically performed just before PET-EEG data acquisition. The assessment with the highest indication for consciousness, or best diagnosis, is retained for final CRS-R diagnosis (Wannez et al., 2017). The consistency of the CRS-R diagnosis corresponds to the percentage of sessions where the patient reaches his/her highest indication for consciousness. In the present dataset, the mean percentage of consistency across patients was 63%±4 SEM. Within the subgroup of patients where FDG-PET analysis indicated MCS but CRS-R indicated UWS (MCS*), the consistency of CRS-R assessments was 100%.

A total of 129 patients were included in the study. As detailed in the Result section ('Extracting of the electrocardiogram from EEG data'), data from 61 patients had to be discarded from further analysis because it was not possible to extract 5 minutes of good quality electrocardiogram from their EEG data. The final analysis of EEG data thus included 68 patients (see Table 1 for demographic information).

The study was approved by the ethics committee of the University Hospital of Liège. Patients' legal guardians gave written informed consent for approval of participation in the study, as required by the declaration of Helsinki.

## Data acquisition

FDG-PET data were acquired as described previously (Stender et al., 2014). In short, the patients fasted for at least 6 hours before commencing the PET procedure. Patients remained in a dark room for approximately 10 minutes before and 30 minutes after injection of 150-300 mBq $^{[18]}$F-FDG, after which the PET was acquired. Patients were in resting state with eyes open. They were aroused when necessary following the same arousal facilitation protocol as used for CRS-R assessment (Giacino et al., 2004).

EEG was recorded during the FDG uptake phase, i.e. after FDG injection, but before tomography begun. High-density EEG recordings were obtained from 256 scalp sensors using saline electrode nets designed by Electric Geodesics, with a sampling rate of 250 or 500Hz. EEG recordings were performed during the dark period of the FDG-PET protocol (i.e. about 10 minutes before until half hour after the $^{[18]}$F-FDG injection). The EEG net was removed before the PET scan was performed. In 8 patients, and 5 additional healthy control subjects, an electrocardiogram (ECG) was obtained from electrodes placed below the left and right clavicles using the Polygraph input box (Electric Geodesics).

## EEG data preprocessing

EEG preprocessing was performed using the Fieldtrip toolbox (Oostenveld et al., 2011) in Matlab R2016b. Data were downsampled from 500 Hz to 250 Hz when applicable, offline filtered (1-25 Hz Butterworth band-pass filter with a Hamming windowing at cutoff frequencies) and z-scores per channel over time were calculated.

The first step in data analysis was to identify a 5-minutes time window of good quality EEG, where chances to derive an ECG from EEG data would be maximized. In each candidate 5-minute time window, we quantified noise as the total duration of EEG segments exceeding a z-scored amplitude of 20, across all channels. In each of the 129 patients, the least noisy 5-minutes time window was retained for further analysis.

Within the selected 5-minutes time window, we then identified artefacted channels. We computed the area under the curve of the z-score amplitude over during the selected 5-minute for each channel and examined the distribution across channels. Channels exceeding +3 standard deviations of the AUC distribution for all channels were discarded (77 ± 4 SEM channels rejected on average, most often located over cheeks and neck). This procedure was iterated until all channels satisfied the +3SD criterion, or more than 50 channels had to be discarded, at which point the participant would be excluded from further analysis, a case that did not occur in the present dataset. To further identify bad channels, we computed the correlation between each channel and its neighbors to identify channels with low correlation, indicative of artefacts such as poor contact. Neighborhood relationships were computed using the default neighborhood definition in Fieldtrip considering as neighbors' channels up to distances of 4 cm. For each channel, we computed a mean correlation score corresponding the mean correlation with neighbors, weighted by the distance between channels. Channels with a

weighted-by-distance correlation lower than 60% were replaced by spline interpolation of neighbors.

After ICA correction (see below for details), we restricted the analysis to the 175 channels located on the scalp, excluding channels on face and neck. Finally the EEG dataset was re-referenced using a common average and reduced to the same 64 channels set in each patient, corresponding to standard scalp locations in the 10-10 system (Luu and Ferree, 2000).

## Extracting the electrocardiogram from EEG data (ICA-ECG)

EEG electrodes measure brain activity but also electrical cardiac activity. This is known as the cardiac artefact (Dirlich et al., 1997). It follows that it is possible to recover the electrocardiogram (ECG) from scalp EEG data (Raimondo et al., 2017). To extract the ECG from EEG, we used Independent Component Analysis (ICA) to obtain ICA-corrected EEG data on the one hand and an electrocardiogram derived from independent component analysis (ICA-ECG) on the other hand. This procedure was successful in 68 patients. Out of 129, 61 patients were discarded at the ICA-EEG stage, 50 because ICA-ECG extraction was not successful, and 11 because artefact-free ICA-ECG could not be obtained in 5 minutes of continuous recordings.

We obtained the independent components using the 'ICA extended' algorithm (Jung et al., 1998). The ICA component showing heart-induced activity with the clearest R peaks and a gradient ear-to-ear amplitude topography (Dirlich et al., 1997) was selected for further analysis. If the R peaks were not clear in part of the 5-minutes segment, we selected a new 5-minute segment, and performed the whole preprocessing again starting from the bad channels' rejection onwards. ICA components related to cardiac activity were removed from EEG signals. The output of this procedure is thus a 5-minute segment of EEG data ICA-corrected for the cardiac

artefact, and a corresponding 5-minute segment of ICA-ECG, i.e. the electrocardiogram extracted by ICA from the EEG data.

R-peaks were detected on the ICA-ECG using an automated process. First, a time window containing at least 5 R-peaks with good signal-to-noise ratio was manually selected. Epochs from 400 ms before to 400 ms after each of those 5 R-peaks were averaged to generate a heartbeat template specific to each patient. The template was then correlated with the entire ECG signal to automatically detect the R-peaks. Correlation's local maxima, indicating the presence of a heartbeat, were detected over the whole 5-minute ECG data, in sliding time windows of a duration equal to the mean interbeat intervals duration obtained in the template plus 200 ms, computed at each sample. To limit the number of false positives, consecutive detected peaks separated by a duration smaller than the estimated mean interbeat interval duration minus 200 ms were discarded automatically. Both peak detection and resulting histogram of interbeat interval duration were visually inspected in each patient and manual addition/removal of peaks were performed if needed (13 ± 2 SEM manual corrections of individual heartbeats on average).

The accuracy of the heartbeat detection procedure from ICA-EEG was subsequently validated in a dataset of 13 participants (8 patients, included in the 68 patients analyzed here, and 5 additional healthy participants) where a regular ECG was also recorded. The validation consisted of a Pearson correlation of interbeat interval time series and power spectrum between 0-0.4 Hz, and percentage error of the power in the three frequency bands typically used to study heart-rate variability (0.03-0.04 Hz, 0.04-0.15, and 0.15-0.4Hz) (Task Force of the European Society of Cardiology the North American Society of Pacing, 1996).

## Heartbeat-evoked responses

The HER corresponds to brain activity evoked by each heartbeat, and can be analyzed by averaging EEG data time-locked to heartbeats (Schandry et al., 1986). The HER was computed after the rejection of cardiac artefacts with independent components analysis, bad channels interpolation and re-referencing to a common average. Epochs of pre-processed EEG data were defined from each R-peak to 500 ms after. Epochs with an amplitude larger than 300 μV on any channel, or where the next or preceding heartbeat occurred at an interval shorter than 500 ms, were discarded. Epoched data were smoothed using a 30 ms time window with an 80% overlap, resulting in 20 time points.

We generated surrogate heartbeats to test whether locking EEG data to real heartbeats, as in the HER analysis, improves classification accuracy as compared to analyzing EEG data at random moments. Heartbeats were reallocated at pseudo-random timings with interbeat intervals larger than 500 ms. In each patient 1000 surrogate heartbeats were generated to compare classification accuracy obtained using random segments of EEG data locked to surrogate heartbeats and the classification accuracy obtained on EEG data locked to physiological heartbeats. Classification based on 500 ms EEG segments locked to surrogate heartbeats was performed using the same rejection criteria and the same classification features as classification based on EEG epochs locked to heartbeats. Note that a number of 500 ms EEG segments necessarily overlap with HERs, since the interval between heartbeats falls most of the time in the 500-1000 ms range.

## Multivariate analysis

We used a machine learning approach of multivariate analysis to distinguish between MCS and UWS patients (Sitt et al., 2014; Raimondo et al., 2017; Engemann et al., 2018). We

followed machine learning good practices to maximally reduce the biases of prediction models (Woo et al., 2017; Steyerberg et al., 2018), such as exploration in a training set with cross-validation followed by prospective validation of the model in an independent test sample, as well as chance level estimation to take into account the influence of an unbalanced number of UWS and MCS patients (Pal, 2005; Verikas et al., 2011).

We first considered a group of 38 patients (training set) for exploration purposes, in which we performed cross-validation to build the prediction model. We performed a 3-fold cross-validation (MCS/UWS in each fold: 11/2, 10/3, 10/2) without hyper-parameter optimization. Its generalization ability was tested in a new, independent set of 30 patients (validation set). Participants were pseudo-randomly assigned to the training set (31 MCS and 7 UWS) and validation set (24 MCS and 6 UWS) to obtain similar proportions of MCS and UWS in both sets by author JA. The authors performing the prospective validation (DCR, CTB) were blind to patients' diagnosis in the validation set.

We used Random Forests (RF) classifier, as implemented at https://code.google.com/archive/p/randomforest-matlab/ (Breiman, 2001), with 1000 trees and a number of features at each node equal to the square root of the total amount of features available, and all other parameters set to default. Compared to other classification approaches, such as Support Vector Machines, RF performs better with noisy and high-dimensional data, limits over-fitting and adapts better to unbalanced datasets (Breiman, 2001; Pal, 2005; Verikas et al., 2011). We nevertheless verified that validation accuracies were larger than chance using a permutation test. We estimated chance level accuracy by computing 1000 classifications where the labels UWS and MCS were randomly assigned, while maintaining the original number of labels in each category as well as the original number of patients in the training and validation sets (Ojala and Garriga, 2010; Combrisson and Jerbi, 2015). We then compared the

distribution of chance level accuracy with the empirical accuracy and derived the corresponding Monte Carlo p-value.

A RF classifier is based on a large number of decision trees that choose their splitting features from a bootstrap sampled subset of features. As a result, one can estimate the relevance of each input feature for classification, using the "Gini impurity index" (Breiman, 2001; Strobl et al., 2008). Additionally, we defined the consciousness score as the proportion of trees that predicted MCS diagnosis. A patient with a consciousness score higher than 0.5 was classified as MCS. If the consciousness score was lower than 0.5 the patient was classified as UWS. Consciousness scores close to 0.5 indicate a more uncertain classification, consciousness scores tending to 1 or 0 indicate that all decision trees reach the same conclusion (MCS or UWS respectively). The consciousness score can be viewed as the classifier confidence about the decision.

## PET data and correlation with consciousness scores

PET data were preprocessed and analyzed to obtain the PET-diagnosis according to (Stender et al., 2014). In short, we used SPM to contrast the glucose metabolism of each patient with the glucose metabolism of a group of 34 healthy control subjects without known neurological or psychological illness. The PET-based diagnosis was made by evaluating the relative preservation of regional metabolism, independently from the CRS-R based diagnosis. In case of complete bilateral hypometabolism of the frontoparietal network cortex the patient was diagnosed as UWS, while partial preservation of the frontoparietal network corresponded to a diagnosis of MCS (Stender et al., 2014).

For further PET analysis, 19 patients were excluded due to the presence of large lesions covering more than 2/3 of an hemisphere and compromising the reliability of anatomical

normalization, as proposed by (Di Perri et al., 2016). A total of 49 patients was thus included in the correlation between cerebral glucose uptake and consciousness score based on HERs. We used Statistical Parametric Mapping 12 (https://www.fil.ion.ucl.ac.uk/spm/), a toolbox developed for Matlab (2017a) to identify brain regions in which glucose metabolism correlated with ECG/EEG-based classification. Separate full factorial designs modeled the effect of the PET-based diagnosis (UWS or MCS) and classifiers' (HER or HRV) consciousness score on cerebral glucose uptake. Proportional scaling was performed to identify regions that showed a relative increase in glucose metabolism with any regressor. T-contrasts were specified for the main effect of the HER/HRV regressors, and for the interaction of the HER/HRV regressor and diagnosis. Results were considered significant at family-wise error (FWE) corrected p-value $<0.05$.

## Heart-rate variability (HRV)

We used standard measurements of HRV as described in (Task Force of the European Society of Cardiology the North American Society of Pacing, 1996), both in the time and in the frequency domains. Standard time domain features included the mean interbeat interval, the standard deviation of the beat-to-beat intervals (SDNN), and the square root of mean squared differences of successive intervals (RMSSD). To estimate power spectral density, we used the Burg periodogram of order 7, a method based on autoregressive spectral estimation which uses parametric methods to model the data (Thayer et al., 1996). Standard frequency domain features for short-term recordings (5 minutes) are defined by the power distribution in three frequency bands: Very low frequency between 0 – 0.04 Hz, low frequency between 0.04 – 0.15 Hz and high frequency between 0.15 – 0.4 Hz.

## Statistical analysis

Statistical comparisons were based on Mann Whitney tests, paired t-tests, chi-square tests or Pearson correlations, as specified in the Results section. Classification accuracy significance was assessed with a Cumulative Binomial Distribution and/or Monte Carlo Permutation Tests. Monte Carlo p-values ($p_{mc}$) were computed as the proportion of the number of classifications (out of the 1000 permutations) in which a higher accuracy was reached, respect the accuracy which significance is being tested. How permutation tests were designed is specified in the corresponding sections of the Material and Methods. Because computing Monte Carlo p-values is computationally intensive, this method could not be used on each of the 1000 draws of random EEG segments. We thus also computed binomial p-values ($p_{bin}$), considering the number of patients as the number of observations, and tested the accuracy significance considering a probability of 0.5, i.e. 1/number of possible diagnosis. To assess binomial significance of the classification on random EEG segments, binomial p-values were computed as the proportion of the number of classifications (out of the 1000 draws of random EEG segments) in which $p_{bin}$ significance was not reached at $\alpha = 0.05$. Bayes factors were computed to estimate evidence for $H_0$ using CRAN package BayesFactor version 0.9.12-4.2 for R as implemented in https://richarddmorey.github.io/BayesFactor/ (Rouder et al., 2009), with default prior r-scale $\sqrt{2}/2$ and substantial evidence for $H_0$ if $BF < 0.3125$ (Kass and Raftery, 1995). The statistical analysis of PET data is described in the PET data section.

## Data and Code Availability Statement

The data supporting this article, that include sensitive health related information, are available upon reasonable request to the Liège Coma Science Group coma@uliege.be

Codes are publicly available at https://github.com/diegocandiar/brain_heart_doc/.

# Results

## Heartbeats detection validation

We first extracted the electrocardiogram from EEG data using independent components analysis (ICA-ECG) (Raimondo et al., 2017). To validate this approach (Fig. 1), we compared ICA-ECG with real ECG in 13 subjects (8 patients and 5 additional healthy controls) where ECG was recorded. Heartbeats were detected independently in ECG and ICA-ECG. The interbeat intervals time series obtained from real ECG and ICA-ECG showed a correlation greater than 0.99 in all 13 subjects (Pearson correlation coefficient, range 0.9928 to 0.9999). Delays of one or two samples between R peaks detected in the real ECG and ICA-ECG could be occasionally observed (Fig. 1C). These minor discrepancies had very limited impact on the spectral analysis of HRV (Pearson correlation coefficient between spectral densities estimated from real ECG and ICA-ECG, range 0.9970 to 1; relative difference between HRV power in very low, low and high-frequency ranges computed using real ECG or ICA-ECG smaller than 5% in all participants).

## Decoding residual consciousness associated with preserved glucose metabolism using heartbeat-evoked responses: exploration in the training sample

We then tested whether HER amplitude and variance could distinguish between UWS and MCS patients in the training sample (n=38). To determine the latencies at which HERs are informative, we computed an independent classifier based on HERs averaged using 200 ms-long sliding time windows and performed a 3-fold cross-validation to estimate classifier performance in each time window. As shown in Fig. 2A, accuracy peaked to 81.62% in the time window between 200 and 400 ms after R peak. To verify that classification results were not driven by cardiac electrical activity (Fig. 2B), we computed the mean ICA-ECG amplitude

in the 200-400 ms time interval, and found that it does not reliably distinguish between MCS (17.72±15.89 SEM) and UWS patients (-3.55±9.11 SEM; Mann Whitney U test, rank sum = 622, z-value = 0.64, p-value = 0.52; BF = 0.44, inconclusive evidence for $H_0$). In addition, we performed RF classification using all the ECG samples of the 200-400 ms time window, and found that accuracy (75.06%±4.51) was much lower than with EEG (81.62%). Because visual inspection suggests a difference in ECG around R peak, we repeated the same analysis for the time window -100 to +100 ms, and again found a lower accuracy (78.32%±3.31) than in EEG. Moreover, the topography of HERs in the 200-400 ms time window (Fig. 2C) was quite different from the typical cardiac artefact distribution (Fig. 1B) (Dirlich et al., 1997).

Having determined the time-window of interest, we proceeded to test whether HER amplitude and variance at each time sample between 200 and 400 ms and at each channel could discriminate between UWS and MCS patients using 3-fold cross-validation. The random forest (RF) classifier had an accuracy of 81.62% in the training sample. The most relevant channels for RF classification (Gini index averaged across amplitude and variance and across time stamps) were located over right temporal and central regions (Fig. 2D). A sustained difference in HERs between UWS and MCS patients could be observed at those locations (Fig. 2E). Classification relied more on HER variance than HER average (paired-sample t-test between Gini impurity indices, t-stat = -3.7411, df = 63, p-value = 0.0004). The analysis of the training sample thus indicates that HERs carry relevant information to detect residual signs of consciousness.

We then tested whether patients' classification using HERs outperforms classification based on EEG segments of the same duration, but not locked to heartbeats. Training classifiers with EEG segments not locked to heartbeats led to a 3-fold cross-validation classification accuracy of 79.08±0.7% in the training sample (average across 1000 classifiers trained on different random selection of EEG segments). Classification accuracy based on random EEG

segments is significantly lower than when EEG is locked to heartbeats (81.62%; Monte Carlo test, $p_{mc}<0.001$; Table 2). Indeed, none of the 1000 classifications performed on random EEG segments had a higher classification accuracy than the classification on HERs, indicating that classification based on HERs is significantly more accurate than classification based on random EEG segments with a $p_{mc}<0.001$. Brain responses to heartbeats thus convey specific additional information on residual consciousness, as compared to generic EEG features.

## Decoding residual consciousness using heartbeat-evoked responses: blind prospective validation in a new sample

Following classification good practices (Woo et al., 2017; Steyerberg et al., 2018), we tested whether such results would generalize to a new sample of 30 participants (blind prospective validation). We trained a classifier using all the folds of the training sample (n=38) and confirmed that the model classifies all patients in the training set with a 100% accuracy before assessing the predictions of this classifier in the validation sample. The classifier reached an accuracy of 86.67% (binomial test, $p_{bin}<0.001$), a sensitivity of 95.8% and specificity of 50% in the validation sample.

We ran two additional control analyses. First, because our sample is unbalanced with more MCS (n=31) than UWS (n=7) patients, we verified that results were not due to chance. To estimate chance level classifier accuracy due to sample imbalance, we computed 1000 classifiers on the same data after randomly reassigning diagnoses across patients, thus maintaining the MCS/UWS imbalance. Chance level accuracy was on average 77.05±3.98%, almost 10 points lower than the accuracy of the classifier trained on real data. Only one random permutation out of 1000 exceeded the accuracy of the classifier trained on the real data, showing that the hypothesis that the 86.67% accuracy of the classifier trained on real data is due to

chance can be rejected with a Monte Carlo $p_{mc}$=0.001. Second, we applied the same procedure to EEG segments not locked to heartbeats and found a mean classification accuracy in the validation sample of 80.84±4.09 % (binomial test over 1000 permutations, $p_{bin}$<0.001), i.e., 6 points below the classification of the classifier based on HERs, a difference that was only marginally significant (comparison between the distribution of the 1000 classifications on random EEG segments with HER-based classifier, $p_{mc}$=0.064). Altogether, the classification results in the validation sample (Table 2) show that HERs provide reliable information about residual signs of consciousness in MCS patients, that can be generalized to a new dataset, and confirm that HERs convey specific consciousness-related information, beyond what can be extracted from EEG data not locked to heartbeats.

## Non-behavioral MCS patients have smaller HER-based consciousness scores

So far, we have considered the diagnosis of consciousness based on brain glucose metabolism (PET-based diagnosis), because this diagnosis is independent from overt behavior. However, clinical diagnosis is most often based on a behavioral index, here based on the best CRS-R subscore, but is not always consistent with the PET-based diagnosis (Gosseries et al., 2014; Stender et al., 2014; Schiff, 2015). In the current dataset, the diagnosis of UWS based on PET data was consistent with the CRS-R diagnosis in all UWS patients. However, patients with a PET-based diagnosis of MCS could be split into two categories: behavioral MCS (n=18), i.e. patients considered as MCS based on both brain metabolism and best CRS-R diagnosis, and non-behavioral MCS (or MCS*, n=6), i.e. patients showing only reflex-like behavior as measured with CRS-R, yet with a brain metabolism suggesting a capacity for consciousness (Gosseries et al., 2014).

For each patient identified as MCS based on PET-diagnosis, we computed a HER-based consciousness score, defined as the proportion of decision trees of the HER-based classifier

predicting MCS diagnosis (Fig. 2F). This score can be viewed as the classifier's confidence in the diagnosis, with four ranges (0-0.25, UWS with high certainty, 0.25-0.5, UWS with low certainty, 0.5-0.75, MCS with low certainty, 0.75-0.1 MCS with high certainty). Patients with consciousness scores indicating MCS with low certainty were significantly more likely to be non-behavioral MCS patients: the proportion of non-behavioral MCS patients was significantly larger in low-certainty MCS patients (55%; 5/9) than in high-certainty MCS patients (6.6%; 1/15; $\chi^2$ test, $\chi^2$-stat = 7.17, p-value = 0.007).

## HER-based consciousness scores correlate with PET metabolism in right occipito-temporal and lateral temporal regions

We next tested whether consciousness-related information conveyed by HERs correlated with glucose metabolism in specific brain regions, in an exploratory analysis. Across both the training and test sets, we identified 49 patients (39 MCS, 10 UWS, based on FDG-PET) whose brain anatomy allowed anatomical normalization, trained a classifier on HERs in those patients, and correlated the resulting HER-based consciousness scores with glucose metabolism. A positive correlation between glucose metabolism and HER-based consciousness scores was found in the right temporal lobe in the 39 MCS patients (Table 3, Fig. 3). The largest cluster extended along the right ventral occipito-temporal regions. Neurosynth (Yarkoni et al., 2011) reveals that this region is functionally connected mostly to its homologous region in the left hemisphere, and that it is strongly associated with the terms "face", "face recognition" and "object recognition", as expected from the known functional organization of the ventral visual system (Mishkin et al., 1983; Haxby et al., 1991; Kanwisher et al., 1997). The second region was located in the anterior part of the right superior temporal sulcus. This region has a much richer pattern of functional connectivity as revealed by Neurosynth, being connected to the core components of the default network, i.e. ventro-medial prefrontal cortex, posterior cingulate

cortex, right inferior parietal lobule, as well as to the right amygdalo-hippocampal region and the homologous region in in the left hemisphere. The lateral temporal region is loosely associated with a larger number of terms, related to social interactions ("social cognitive", "theory of mind", "sentence comprehension") but also to internal, self-related cognition ("default network", "autobiographical"). No significant correlation could be found between HER-based consciousness score and PET metabolism in the 10 UWS patients.

## Using HERs to decode consciousness as diagnosed from CRS-R

In this study we explore markers of consciousness independent from behavior. However, the measure of consciousness most commonly used in clinical practice is the behavioral assessments with the CRS-R. We thus tested whether we could decode consciousness as diagnosed with CRS-R using HERs, using the same procedure as used for decoding consciousness as diagnosed with PET, and compared CRS-R-based classification accuracy to PET-based classification accuracy.

To determine the latencies at which HERs are informative, we computed an independent classifier based on HERs averaged using 200 ms-long sliding time windows and performed a 3-fold cross-validation to estimate classifier performance in each time window. The best candidate time-window to decode consciousness as diagnosed with CRS-R in the training sample was 156-356 ms, i.e. quite close to the best time window to decode consciousness from PET diagnosis, which was 200-400 ms. We then trained a classifier with HER amplitude and variance at each channel and each time sample of the 156-356 ms time-window using 3-fold cross-validation. The accuracy of classification based on CRS-R diagnosis was 79.06% (binomial test, $p_{bin}<0.001$), which is somewhat lower than the classification based on PET diagnosis (3-fold cross validation accuracy 81.62%). HERs did convey a significant advantage to CRS-R based classification: classifications performed on random EEG segments led to an

average accuracy of 72.95±4.49% (binomial test over 1000 permutations, $p_{bin}$=0.012), whereas HER-based classification accuracy was 79.06% (comparison between the accuracies distribution of the 1000 classifications on random EEG segments and the HER-based classifier accuracy, Monte Carlo $p_{mc}$=0.038).

We then trained a CRS-R-based classifier on all the folds of the training sample (n=38) and tested the performance of this classifier in the validation set (n=30, blind prospective cross-validation). The classification performed over CRS-R diagnosis reached an accuracy of 76.67%, which is significantly above chance (chance accuracy obtained by randomly shuffling diagnosis labels, 59.53±5.34%; 4 out of 1000 permutations outperformed classification on real diagnosis, Monte Carlo $p_{mc}$=0.004). HERs did convey a significant advantage to CRS-R-based classification, as compared to EEG segments not locked to heartbeats (EEG segments, mean classification accuracy 66,42±7.13% (binomial test over 1000 permutations, $p_{bin}$=0.272), HER-based classification accuracy, 76.67%, binomial test, $p_{bin}$<0.001, Monte Carlo test, $p_{mc}$=0.041). However, the blind prospective classification based on CRS-R diagnosis remained 10 points lower than the accuracy of the classifier based on PET diagnosis which was 86.67% (Table 2). Finally, when the classification used CRS-R diagnosis, HER consciousness scores did not correlate with brain glucose metabolism, in any region. HER classification using the CRS-R diagnosis of consciousness thus reproduces partly the results obtained with HER classification using the PET diagnosis of consciousness, such as a better accuracy of HERs than random EEG segments. However, the generalization accuracy of the CRS-R-based HER classifier was much lower than the generalization accuracy the PET-based HER classifier. This suggests an underlying better consistency between two brain measures of consciousness (PET glucose metabolism and HERs) than between a behavioral measure of consciousness (CRS-R) and a neural measure (HER). Of note, in the present study, the difference between CRS-R diagnosis and PET diagnosis is exclusively found in MCS* patients, where PET indicates minimal

consciousness while CRS-R indicates unconsciousness. All 5 behavioral assessments using CRS-R, including the one performed at the time of EEG-PET data acquisition, indicated unconsciousness in all MCS* patients with 100% consistency over time. The difference between PET-based diagnosis and CRS-R diagnosis thus seems to correspond to a reliable difference between methods of assessments rather than to fluctuations in behavior.

## No reliable decoding of residual consciousness using heart-rate variability

Our results show that HERs, reflecting how the brain responds to ascending signal at each heartbeat, could accurately distinguish MCS from UWS patients, especially when using PET-based consciousness diagnosis, which is independent from overt behavioral responses. We further tested whether heart rate variability, which is strongly constrained by descending signals from brain to heart, can distinguish between MCS and UWS during resting state. We first investigated, using a standard univariate approach (Mann Whitney U-tests between MCS and UWS), whether routinely used measures of HRV would differ between MCS and UWS patients in the training set. None of the classical measures (Task Force of the European Society of Cardiology the North American Society of Pacing, 1996) we computed could reliably distinguish between MCS and UWS patients (Table 4).

Although neither the low nor high frequency power of HRV could distinguish between MCS and UWS, the HRV power spectrum density showed qualitative differences (Fig. 4A), suggesting that the full power spectrum is a good candidate for multi-variate analysis. Using HRV frequency power at 13 different frequencies as features, we obtained a 3-fold cross-validation accuracy of 79.06% in the training sample. Two frequency ranges were particularly relevant for classification (Fig. 4B), ~0.1-0.2 Hz and ~0.3-0.4 Hz, that differ slightly from the classical low (0.04-0.15 Hz) and high frequency (0.15-0.4 Hz) HRV ranges (Task Force of the

European Society of Cardiology the North American Society of Pacing, 1996). We then trained a classifier using all folds from the training set, reaching a 100% accuracy. We tested this classifier in the independent validation sample of 30 patients, which resulted in an accuracy of 76.67%. HRV validation accuracy was not significantly different than chance accuracy (Monte Carlo test, $p_{mc} = 0.382$) and was 10 points lower than HER validation accuracy (86.67%).

The classification based on HER and the classification based on HRV generated different predictions, and misclassified patients with one or the other method were different (Fig. 4C). Consciousness scores obtained from the HRV-based classifier did not correlate with the consciousness scores obtained from the HER-based classifier (Pearson correlation r = –0.1512, p-value = 0.4251, BF = 0.5241, inconclusive evidence for $H_0$). These results suggest that HER-based and HRV-based classification might be complementary. We thus tried to combine predictions based on HER and HRV, by applying the following rule: a patient is classified MCS if both the HER-based and HRV-based consciousness scores are larger 0.5, else the patient is classified as UWS. This procedure led to an accuracy of 83.33% (binomial test, $p_{bin}<0.001$) in the validation sample, i.e. a lower accuracy than with HER-based classification only (86.67%). However, combining HER and HRV-based classification led to a better balance between sensitivity and specificity, compared to HER only.

## Discussion

Because ascending signals from heart to brain have been associated with different aspects of consciousness in healthy participants (Park et al., 2014, 2016; Babo-Rebelo et al., 2016a, 2016b, 2019; Sel et al., 2017), we hypothesized that neural responses to heartbeats would differ in MCS compared to UWS patients. We found that heartbeat-evoked responses could accurately distinguish MCS and UWS patients in an independent validation sample with an accuracy of 87%, (sensitivity 96%; specificity 50%), when the consciousness diagnosis was based on glucose uptake as measured with PET, i.e., a diagnosis independent from behavioral output. Using HERs as features systematically led to classifier accuracies that were larger than when using on random EEG segments or heart-rate variability as features, showing that how the brain responds to heartbeats provides additional information on consciousness, as compared to generic ongoing EEG activity or to heart rate. HER-based classification offered not only a dichotomous classification between UWS and MCS patients but also a more graded and individualized consciousness score which captured some of the complexity of the clinical diagnosis, with low consciousness scores being much more frequent in non-behavioral MCS. HER-based consciousness scores correlated with glucose metabolism in the default mode network node located in the right superior temporal sulcus, as well as with the right ventral occipito-temporal cortex. HERs predicted consciousness better when consciousness was inferred from glucose metabolism than when consciousness was inferred from behavior using CRS-R diagnosis. Our results indicate that the transient brain responses elicited at each heartbeat convey relevant information on residual consciousness that does not necessarily translate into behavior, in patients with chronic disorders of consciousness.

## HERs improve the detection of residual consciousness

HERs correspond to brain responses to ascending cardiac inputs at each heartbeat. So far, cardiac-related indexes of consciousness have been obtained from HRV, that characterizes descending control of the brain on cardiac rhythm. HRV distinguishes between UWS and MCS patients receiving nociceptive stimulation (Leo et al., 2016; Riganello et al., 2018a; Tobaldini et al., 2018), auditory stimulation (Raimondo et al., 2017), or at rest and under sedation (Riganello et al., 2018b). Here, we show that HERs are more efficient to distinguish between UWS and MCS patients in resting, awake state than HRV. Several lines of evidence suggest that the difference in HERs between MCS and UWS is of cerebral origin, rather than cardiac electrical activity detected at the sensor level. First, we could not find any ECG amplitude difference between MCS and UWS patients, indicating that the HER effect is not directly driven by cardiac electrical activity. Second, the most relevant channels to distinguish between MCS and UWS patients were located at right central sites, while the EEG channels most affected by the cardiac artefact display the typical topography over left temporal and right frontal regions (Dirlich et al., 1997). Third, HER and HRV provided non redundant information on consciousness. Last, HER-based consciousness scores correlated with brain glucose uptake measured with PET in right temporal regions, suggesting that HER-based classification is directly related to brain activity.

A number of studies have successfully used EEG to predict (un)consciousness, with prospective validation accuracies in the 71-78% range (Engemann et al., 2018). Our own results when using random EEG segments not locked to heartbeats (prospective validation accuracy 81%) commensurate with the literature, and confirm that generic EEG features, as captured by random EEG segments not locked to heartbeats, do provide valuable information to detect residual consciousness. However, the classifications we performed on HERs, using both amplitude (Raimondo et al., 2017) and variance from one trial to the next as features, were

systematically more accurate than those performed on random EEG segments. HERs thus convey reliable and specific information about residual consciousness in addition to that provided by the generic EEG features captured by random EEG segments.

## HERs as an indicator of conscious inner mental life even in the absence of behavior

The hypothesis that motivated this work is that HERs contribute to subjective experience (Park and Tallon-Baudry, 2014; Tallon-Baudry et al., 2018; Azzalini et al., 2019). While the results are in keeping with this hypothesis, alternative interpretations have to be considered.

HERs might be a marker of an overall brain state fostering consciousness, such as arousal. However, an overall brain state should also be reflected in neural activity not locked to heartbeats, whereas we found a systematic advantage to classifications based on HERs over classifications based on random EEG segments. Besides, the correlations between glucose uptake and HER-based consciousness scores did not occur in brain regions associated with arousal, such as the saliency network. Last, in healthy participants, consciousness-related modulations of HERs were not accompanied by changes in arousal (Park et al., 2014; Babo-Rebelo et al., 2016a, 2019). The difference between classifications based on random EEG segments and segments aligned to cardiac events is thus likely to be attributed to a change in EEG amplitude and/or variance occurring specifically after heartbeats, i.e. the definition of a heartbeat-evoked response.

If HERs do not reflect a global brain state, then what do they reflect? HERs might convey consciousness-related information about how the brain responds to an internal stimulus, much as responses to exteroceptive stimuli (Cruse et al., 2011; Sitt et al., 2014) or artificial stimuli such as transcranial magnetic pulses (Casali et al., 2013; Casarotto et al., 2016) are

informative about consciousness state. However, we did not find any link between HER-based consciousness scores and glucose uptake in interoceptive regions such as primary somatosensory cortex (Kern et al., 2013) or insula (Park et al., 2018). The present results relating HERs to residual consciousness in patients appear more congruent with the covariation between HERs and perceptual, bodily and self-consciousness (Park et al., 2014, 2016; Babo-Rebelo et al., 2016a, 2016b, 2019; Sel et al., 2017; Al et al., 2020) observed in healthy participants.

We could link HER-based classification results to glucose metabolism in two regions, that are very different from the fronto-parietal regions used to determine the diagnosis of consciousness based on glucose metabolism. The right occipito-temporal cluster corresponds to areas involved in visual shape analysis, while the right anterior superior temporal sulcus is known for its role for in spontaneous, potentially self-related, cognition (Andrews-Hanna et al., 2010), as well as in self-recognition and theory of mind (van Veluw and Chance, 2014). The right lateralization of the brain regions whose glucose metabolism correlates with HER-based classification scores, as well as of the electrodes contributing to the HER-based classification, is reminiscent of non-verbal consciousness in the right hemisphere of split-brain patients (Sperry, 1984).

Consciousness is defined by the existence of subjective experience and inner mental life, a feature that does not necessarily translate into overt behavior. HERs appear here as a valid indicator of inner mental life, since HER-based classification successfully detects residual consciousness even during resting-state without behavioral reports, and even in non-behavioral MCS patients, who never showed any behavioral sign of consciousness. HERs might thus help specify the gray zone of consciousness, i.e., the fleeting conscious feelings that might not be cognitively accessed nor translate into behavioral outputs. This interpretation is in keeping with the fact the HER-based classification is in better agreement with the consciousness diagnosis

based on brain metabolism than with the consciousness diagnosis based on behavior. It is worth noting that this better agreement does not reflect a mere correlation between two different measures of brain activity, independently of the notion of consciousness, since the link between HERs and glucose metabolism is mediated by the consciousness diagnosis on which the HER classifier is trained. Because measuring consciousness in the absence of overt behavior is challenging, cross-validating measures of consciousness is an important step to detect consciousness at the cerebral, rather than at the behavioral level, be it at bedside (Bodart et al., 2017) or in healthy participants (Tsuchiya et al., 2015; Bayne et al., 2017).

## Potential clinical relevance

Observations that heartbeat evoked responses are linked to consciousness in healthy adults (Park et al., 2014, 2016; Babo-Rebelo et al., 2016a, 2016b, 2019; Sel et al., 2017) could lead to a new sensitive tool to identify residual consciousness and establish fine-grained stratification in patients with disorders of consciousness, with only 5 minutes of resting-state EEG data that can easily be acquired at bedside, that could prove particularly valuable when PET, fMRI or TMS-EEG are not available or cannot be employed for safety reasons. Note that recording an ECG together with EEG would markedly improve the procedure, since the extraction of ECG from EEG using ICA is not always possible. Because the experimental situation does not require the active involvement of the patient beyond being awake, the patient's specific sensory impairments or cognitive deficits should have a limited impact on HER-based classification. Still, the clinical relevance of neural responses to heartbeats remains to be validated in a larger, multi-site cohort, with more UWS patients, and patients long term outcome. When assessing consciousness along multiple dimensions (Sitt et al., 2014; Bayne et al., 2017; Sergent et al., 2017; Song et al., 2018), neural responses to heartbeats could contribute to the evaluation of more experiential aspects of consciousness, potentially corresponding to

fleeting conscious feelings not accompanied by behavioral signs of consciousness and not detected by standard clinical assessment based on CRS diagnosis, rather than fully developed intentional communication that translates into behavior.

## Conclusion

HERs convey specific information on residual consciousness in patients, as compared to EEG not locked to heartbeats. This result lends support to the hypothesis that HERs play a role in generating conscious experience (Park and Tallon-Baudry, 2014; Tallon-Baudry et al., 2018; Azzalini et al., 2019), and complement a number of experimental results in healthy participants pointing in the same direction (for review, (Azzalini et al., 2019)). We further showed that HERs capture the consciousness diagnosis based on brain metabolism during resting state better than the consciousness diagnosis based on behavior. It is thus possible to conceptualize and experimentally test consciousness as an experiential phenomenon, rather than as an intermediate cognitive step between external input and behavioral output.

## Author Contributions

D. C.-R., J. A., S. L. and C. T.-B. designed the study; J. A., O.G., C.M., A. T. and S.L. acquired data and established clinical diagnoses; D. C.-R., J. A., C. T.-B. analyzed data; D. C.-R., J. A. and C. T.-B. wrote the initial version of manuscript. All authors contributed to the submitted version.

## References


Al E, Iliopoulos F, Forschack N, Nierhaus T, Grund M, Motyka P, Gaebler M, Nikulin VV, Villringer A (2020) Heart–brain interactions shape somatosensory perception and evoked potentials. PNAS 117:10575–10584.

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL (2010) Functional-anatomic fractionation of the brain's default network. Neuron 65:550–562.

Azzalini D, Rebollo I, Tallon-Baudry C (2019) Visceral Signals Shape Brain Dynamics and Cognition. Trends in Cognitive Sciences 23:488–509.

Babo-Rebelo M, Buot A, Tallon-Baudry C (2019) Neural responses to heartbeats distinguish self from other during imagination. NeuroImage 191:10–20.

Babo-Rebelo M, Richter CG, Tallon-Baudry C (2016a) Neural Responses to Heartbeats in the Default Network Encode the Self in Spontaneous Thoughts. J Neurosci 36:7829–7840.

Babo-Rebelo M, Wolpert N, Adam C, Hasboun D, Tallon-Baudry C (2016b) Is the cardiac monitoring function related to the self in both the default network and right anterior insula? Philos Trans R Soc Lond, B, Biol Sci 371.

Bayne T, Hohwy J, Owen AM (2017) Reforming the taxonomy in disorders of consciousness. Annals of Neurology 82:866–872.

Blanke O, Metzinger T (2009) Full-body illusions and minimal phenomenal selfhood. Trends Cogn Sci (Regul Ed) 13:7–13.

Block N (2005) Two neural correlates of consciousness. Trends Cogn Sci (Regul Ed) 9:46–52.

Bodart O, Gosseries O, Wannez S, Thibaut A, Annen J, Boly M, Rosanova M, Casali AG, Casarotto S, Tononi G, Massimini M, Laureys S (2017) Measures of metabolism and complexity in the brain of patients with disorders of consciousness. NeuroImage: Clinical 14:354–362.

Breiman L (2001) Random Forests. Machine Learning 45:5–32.

Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Casali KR, Casarotto S, Bruno M-A, Laureys S, Tononi G, Massimini M (2013) A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior. Science Translational Medicine 5:198ra105-198ra105.

Casarotto S, Comanducci A, Rosanova M, Sarasso S, Fecchio M, Napolitani M, Pigorini A, G Casali A, Trimarchi PD, Boly M, Gosseries O, Bodart O, Curto F, Landi C, Mariotti M, Devalle G, Laureys S, Tononi G, Massimini M (2016) Stratification of unresponsive patients by an independently validated index of brain complexity. Ann Neurol 80:718–729.

Chalmers DJ (1995) Facing Up to the Problem of Consciousness. Journal of Consciousness Studies 2:200–19.

Combrisson E, Jerbi K (2015) Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. J Neurosci Methods 250:126–136.

Cruse D, Chennu S, Chatelle C, Bekinschtein TA, Fernández-Espejo D, Pickard JD, Laureys S, Owen AM (2011) Bedside detection of awareness in the vegetative state: a cohort study. Lancet 378:2088–2094.

Di Perri C, Bahri MA, Amico E, Thibaut A, Heine L, Antonopoulos G, Charland-Verville V, Wannez S, Gomez F, Hustinx R, Tshibanda L, Demertzi A, Soddu A, Laureys S (2016) Neural correlates of consciousness in patients who have emerged from a minimally conscious state: a cross-sectional multimodal imaging study. The Lancet Neurology 15:830–842.

Dirlich G, Vogl L, Plaschke M, Strian F (1997) Cardiac field effects on the EEG. Electroencephalography and Clinical Neurophysiology 102:307–315.

Engemann DA, Raimondo F, King J-R, Rohaut B, Louppe G, Faugeras F, Annen J, Cassol H, Gosseries O, Fernandez-Slezak D, Laureys S, Naccache L, Dehaene S, Sitt JD (2018) Robust EEG-based cross-site and cross-protocol classification of states of consciousness. Brain 141:3179–3192.

Frässle S, Sommer J, Jansen A, Naber M, Einhäuser W (2014) Binocular Rivalry: Frontal Activity Relates to Introspection and Action But Not to Perception. J Neurosci 34:1738–1747.

Giacino JT, Ashwal S, Childs N, Cranford R, Jennett B, Katz DI, Kelly JP, Rosenberg JH, Whyte J, Zafonte RD, Zasler ND (2002) The minimally conscious state: definition and diagnostic criteria. Neurology 58:349–353.

Giacino JT, Kalmar K, Whyte J (2004) The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility. Arch Phys Med Rehabil 85:2020–2029.

Giacino JT, Katz DI, Schiff ND, Whyte J, Ashman EJ, Ashwal S, Barbano R, Hammond FM, Laureys S, Ling GSF, Nakase-Richardson R, Seel RT, Yablon S, Getchius TSD, Gronseth GS, Armstrong MJ (2018) Practice Guideline Update Recommendations Summary: Disorders of Consciousness: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology; the American Congress of Rehabilitation Medicine; and the National Institute on Disability, Independent Living, and Rehabilitation Research. Archives of Physical Medicine and Rehabilitation 99:1699–1709.

Gosseries O, Zasler ND, Laureys S (2014) Recent advances in disorders of consciousness: Focus on the diagnosis. Brain Injury 28:1141–1150.

Haxby JV, Grady CL, Horwitz B, Ungerleider LG, Mishkin M, Carson RE, Herscovitch P, Schapiro MB, Rapoport SI (1991) Dissociation of object and spatial visual processing pathways in human extrastriate cortex. PNAS 88:1621–1625.

Jung T-P, Humphries C, Lee T-W, Makeig S, McKeown MJ, Iragui V, Sejnowski TJ (1998) Extended ICA Removes Artifacts from Electroencephalographic Recordings. In:

Advances in Neural Information Processing Systems 10 (Jordan MI, Kearns MJ, Solla SA, eds), pp 894–900. MIT Press.

Kanwisher N, McDermott J, Chun MM (1997) The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. J Neurosci 17:4302–4311.

Kass RE, Raftery AE (1995) Bayes Factors. Journal of the American Statistical Association 90:773–795.

Kern M, Aertsen A, Schulze-Bonhage A, Ball T (2013) Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. Neuroimage 81:178–190.

Kondziella D, Bender A, Diserens K, Erp W van, Estraneo A, Formisano R, Laureys S, Naccache L, Ozturk S, Rohaut B, Sitt JD, Stender J, Tiainen M, Rossetti AO, Gosseries O, Chatelle C (2020) European Academy of Neurology guideline on the diagnosis of coma and other disorders of consciousness. European Journal of Neurology 27:741–756.

Laureys S, Celesia GG, Cohadon F, Lavrijsen J, León-Carrión J, Sannita WG, Sazbon L, Schmutzhard E, von Wild KR, Zeman A, Dolce G, the European Task Force on Disorders of Consciousness (2010) Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. BMC Medicine 8:68.

Laureys S, Owen AM, Schiff ND (2004) Brain function in coma, vegetative state, and related disorders. The Lancet Neurology 3:537–546.

Leo A, Naro A, Cannavò A, Pisani LR, Bruno R, Salviera C, Bramanti P, Calabrò RS (2016) Could autonomic system assessment be helpful in disorders of consciousness diagnosis? A neurophysiological study. Exp Brain Res 234:2189–2199.

Luu P, Ferree TC (2000) Determination of the Geodesic Sensor Nets' Average Electrode Positions and Their 10 – 10 International Equivalents. Available at: https://www.egi.com/images/HydroCelGSN_10-10.pdf [Accessed April 6, 2018].

Mishkin M, Ungerleider LG, Macko KA (1983) Object vision and spatial vision: two cortical pathways. Trends in Neurosciences 6:414–417.

Nakayama N, Okumura A, Shinoda J, Nakashima T, Iwama T (2006) Relationship between regional cerebral metabolism and consciousness disturbance in traumatic diffuse brain injury without large focal lesions: an FDG-PET study with statistical parametric mapping analysis. J Neurol Neurosurg Psychiatry 77:856–862.

Ojala M, Garriga GC (2010) Permutation Tests for Studying Classifier Performance. Journal of Machine Learning Research 11:1833–1863.

Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience 2011:9 pages.

Pal M (2005) Random forest classifier for remote sensing classification. International Journal of Remote Sensing 26:217–222.

Park H-D, Bernasconi F, Bello-Ruiz J, Pfeiffer C, Salomon R, Blanke O (2016) Transient Modulations of Neural Responses to Heartbeats Covary with Bodily Self-Consciousness. J Neurosci 36:8453–8460.

Park H-D, Bernasconi F, Salomon R, Tallon-Baudry C, Spinelli L, Seeck M, Schaller K, Blanke O (2018) Neural Sources and Underlying Mechanisms of Neural Responses to Heartbeats, and their Role in Bodily Self-consciousness: An Intracranial EEG Study. Cereb Cortex 28:2351–2364.

Park H-D, Correia S, Ducorps A, Tallon-Baudry C (2014) Spontaneous fluctuations in neural responses to heartbeats predict visual detection. Nat Neurosci 17:612–618.

Park H-D, Tallon-Baudry C (2014) The neural subjective frame: from bodily signals to perceptual consciousness. Phil Trans R Soc B 369:20130208.

Raimondo F, Rohaut B, Demertzi A, Valente M, Engemann DA, Salti M, Fernandez Slezak D, Naccache L, Sitt JD (2017) Brain–heart interactions reveal consciousness in noncommunicating patients. Annals of Neurology 82:578–591.

Riganello F, Chatelle C, Schnakers C, Laureys S (2018a) Heart Rate Variability as an indicator of nociceptive pain in disorders of consciousness? J Pain Symptom Manage 57:47–56.

Riganello F, Larroque SK, Bahri MA, Heine L, Martial C, Carrière M, Charland-Verville V, Aubinet C, Vanhaudenhuyse A, Chatelle C, Laureys S, Di Perri C (2018b) A Heartbeat Away From Consciousness: Heart Rate Variability Entropy Can Discriminate Disorders of Consciousness and Is Correlated With Resting-State fMRI Brain Connectivity of the Central Autonomic Network. Front Neurol 9.

Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G (2009) Bayesian t tests for accepting and rejecting the null hypothesis. Psychonomic Bulletin & Review 16:225–237.

Rudolf J, Ghaemi M, Ghaemi M, Haupt WF, Szelies B, Heiss WD (1999) Cerebral glucose metabolism in acute and persistent vegetative state. J Neurosurg Anesthesiol 11:17–24.

Schandry R, Sparrer B, Weitkunat R (1986) From the heart to the brain: a study of heartbeat contingent scalp potentials. Int J Neurosci 30:261–275.

Schiff ND (2015) Cognitive Motor Dissociation Following Severe Brain Injuries. JAMA Neurol 72:1413–1415.

Sel A, Azevedo RT, Tsakiris M (2017) Heartfelt Self: Cardio-Visual Integration Affects Self-Face Recognition and Interoceptive Cortical Processing. Cereb Cortex 27:5144–5155.

Sergent C, Faugeras F, Rohaut B, Perrin F, Valente M, Tallon-Baudry C, Cohen L, Naccache L (2017) Multidimensional cognitive evaluation of patients with disorders of consciousness using EEG: A proof of concept study. NeuroImage: Clinical 13:455–469.

Sitt JD, King J-R, El Karoui I, Rohaut B, Faugeras F, Gramfort A, Cohen L, Sigman M, Dehaene S, Naccache L (2014) Large scale screening of neural signatures of

consciousness in patients in a vegetative or minimally conscious state. Brain 137:2258–2270.

Song M, Yang Y, He J, Yang Z, Yu S, Xie Q, Xia X, Dang Y, Zhang Q, Wu X, Cui Y, Hou B, Yu R, Xu R, Jiang T (2018) Prognostication of chronic disorders of consciousness using brain functional networks and clinical characteristics Stephan KE, Behrens TE, eds. eLife 7:e36173.

Sperry R (1984) Consciousness, personal identity and the divided brain. Neuropsychologia 22:661–673.

Stender J, Gosseries O, Bruno M-A, Charland-Verville V, Vanhaudenhuyse A, Demertzi A, Chatelle C, Thonnard M, Thibaut A, Heine L, Soddu A, Boly M, Schnakers C, Gjedde A, Laureys S (2014) Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study. The Lancet 384:514–522.

Stender J, Kupers R, Rodell A, Thibaut A, Chatelle C, Bruno M-A, Gejl M, Bernard C, Hustinx R, Laureys S, Gjedde A (2015) Quantitative Rates of Brain Glucose Metabolism Distinguish Minimally Conscious from Vegetative State Patients. J Cereb Blood Flow Metab 35:58–65.

Steyerberg EW, Uno H, Ioannidis JPA, Calster B van, Ukaegbu C, Dhingra T, Syngal S, Kastrinos F (2018) Poor performance of clinical prediction models: the harm of commonly applied methods. Journal of Clinical Epidemiology 98:133–143.

Strobl C, Boulesteix A-L, Kneib T, Augustin T, Zeileis A (2008) Conditional variable importance for random forests. BMC Bioinformatics 9:307.

Tallon-Baudry C, Campana F, Park H-D, Babo-Rebelo M (2018) The neural monitoring of visceral inputs, rather than attention, accounts for first-person perspective in conscious vision. Cortex 102:139–149.

Task Force of the European Society of Cardiology the North American Society of Pacing (1996) Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use. Circulation 93:1043–1065.

Thayer JF, Peasley C, Muth ER (1996) Estimation of respiratory frequency from autoregressive spectral analysis of heart period. Biomed Sci Instrum 32:93–99.

Thibaut A, Bruno M-A, Chatelle C, Gosseries O, Vanhaudenhuyse A, Demertzi A, Schnakers C, Thonnard M, Charland-Verville V, Bernard C, Bahri M, Phillips C, Boly M, Hustinx R, Laureys S (2012) Metabolic activity in external and internal awareness networks in severely brain-damaged patients. J Rehabil Med 44:487–494.

Tobaldini E, Toschi-Dias E, Trimarchi PD, Brena N, Comanducci A, Casarotto S, Montano N, Devalle G (2018) Cardiac autonomic responses to nociceptive stimuli in patients with chronic disorders of consciousness. Clinical Neurophysiology 129:1083–1089.

Tommasino C, Grana C, Lucignani G, Torri G, Fazio F (1995) Regional cerebral metabolism of glucose in comatose and vegetative state patients. J Neurosurg Anesthesiol 7:109–116.

Tsuchiya N, Wilke M, Frässle S, Lamme VAF (2015) No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. Trends Cogn Sci (Regul Ed) 19:757–770.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. NeuroImage 15:273–289.

van Veluw SJ, Chance SA (2014) Differentiating between self and others: an ALE meta-analysis of fMRI studies of self-recognition and theory of mind. Brain Imaging and Behavior 8:24–38.

Verikas A, Gelzinis A, Bacauskiene M (2011) Mining data with random forests: A survey and results of new tests. Pattern Recognition 44:330–349.

Wannez S, Heine L, Thonnard M, Gosseries O, Laureys S, Coma Science Group collaborators (2017) The repetition of behavioral assessments in diagnosis of disorders of consciousness. Ann Neurol 81:883–889.

Woo C-W, Chang LJ, Lindquist MA, Wager TD (2017) Building better biomarkers: brain models in translational neuroimaging. Nat Neurosci 20:365–377.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. Nature Methods 8:665–670.

## Figure legends.

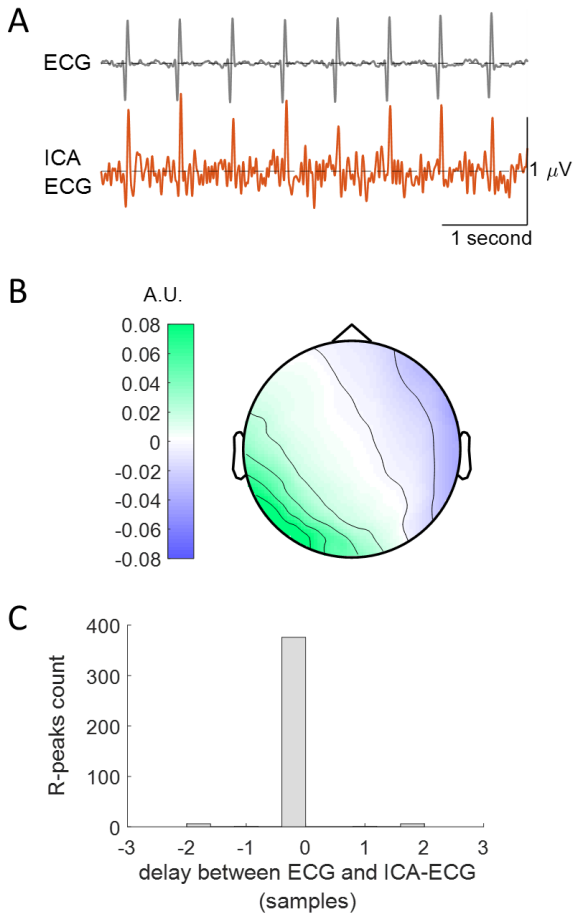


**Figure 1. Extraction of the electrocardiogram from EEG data, example in one patient where both EEG and ECG were acquired.** (**A**) Real electrocardiogram acquired from chest electrodes (ECG, top) and electrocardiogram extracted from EEG data (ICA-ECG, bottom). (**B**) Typical topography of the cardiac-related ICA component, corresponding to the cardiac artefact in EEG data. (**C**) Histogram of delays between R peaks measured in ECG and ICA-ECG. Within the 390 heartbeats recorded in 5 minutes, most R-peaks were perfectly aligned and only 14 (3.6%) R peaks showed a delay of 1 or 2 samples (2 to 4 ms).

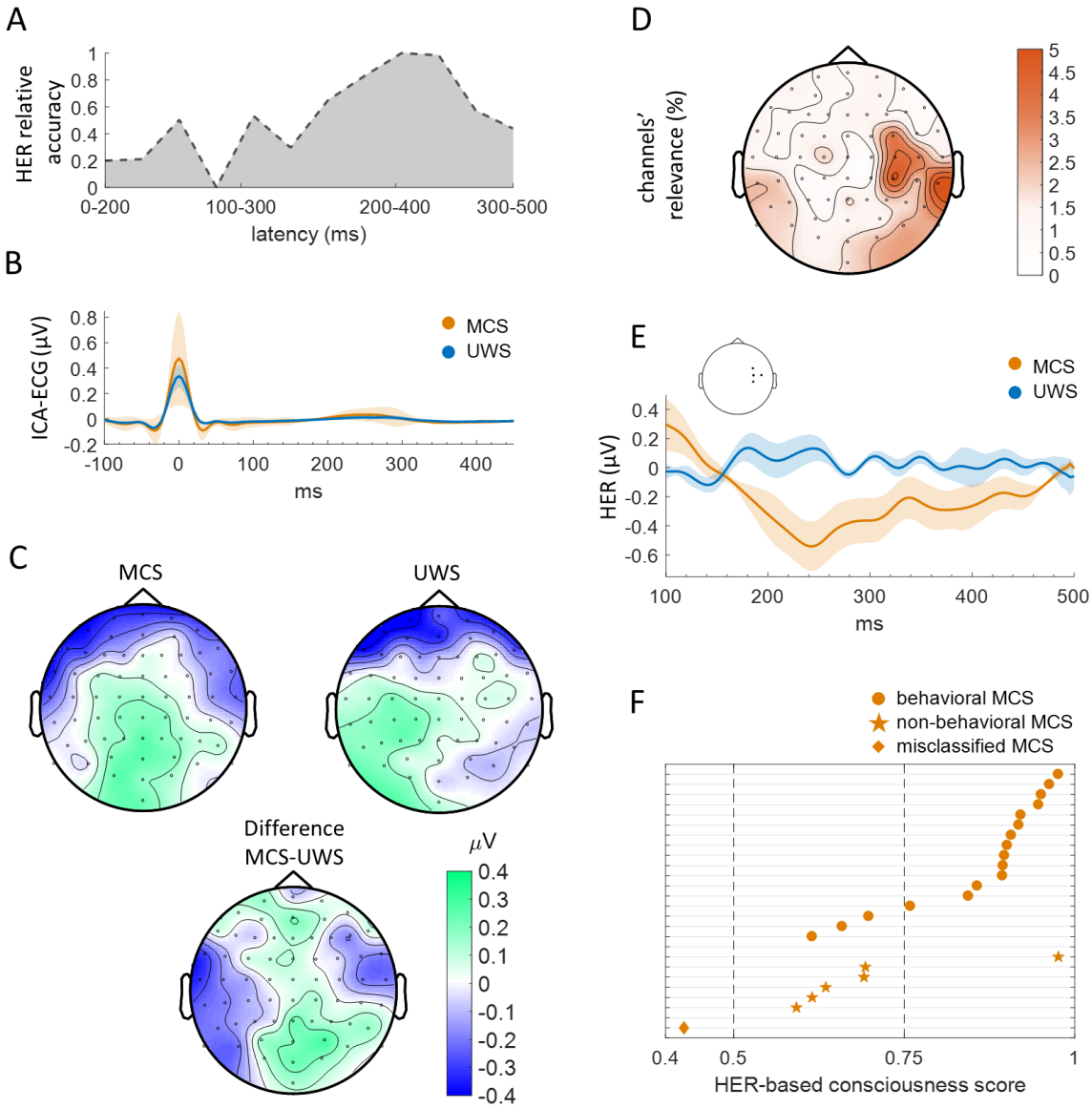


**Figure 2. Classification of PET-based diagnosis of consciousness based on heartbeat-evoked responses.**

(**A**) Classification accuracy in sliding 200 ms-long time windows in a 3-fold cross-validation with HER amplitude and variance at each channel as features, using the consciousness diagnosis obtained from brain glucose metabolism measure with FDG positron emission tomography (PET), in the training sample. Accuracy peaks in the 200-400 ms window. Accuracies have been normalized (posterior min-max normalization to unitary variance), 1 corresponding to best accuracy across all time windows. (**B**) ICA-ECG in MCS and UWS

patients do not differ between 200 and 400 ms. (**C**) HER topographies averaged between 200 and 400 ms, for MCS patients, UWS patients and between-group difference, in the training sample. The topographies markedly differ from the cardiac artefact (Fig. 1B) and indicate the neural origin of HERs. (**D**) Topography of channel relevance in the training sample, based on both amplitude and variance of HERs, showing a larger contribution of left and right occipito-temporal electrodes as well as of right central and temporal electrodes. (**E**) Group-average HERs in the training sample, averaged across right central channels, showing a sustained difference between 200 and 400 ms. (**F**) HER-based consciousness scores in the independent validation sample, showing low HER-based consciousness scores are more likely to correspond to non-behavioral MCS than to behavioral MCS.

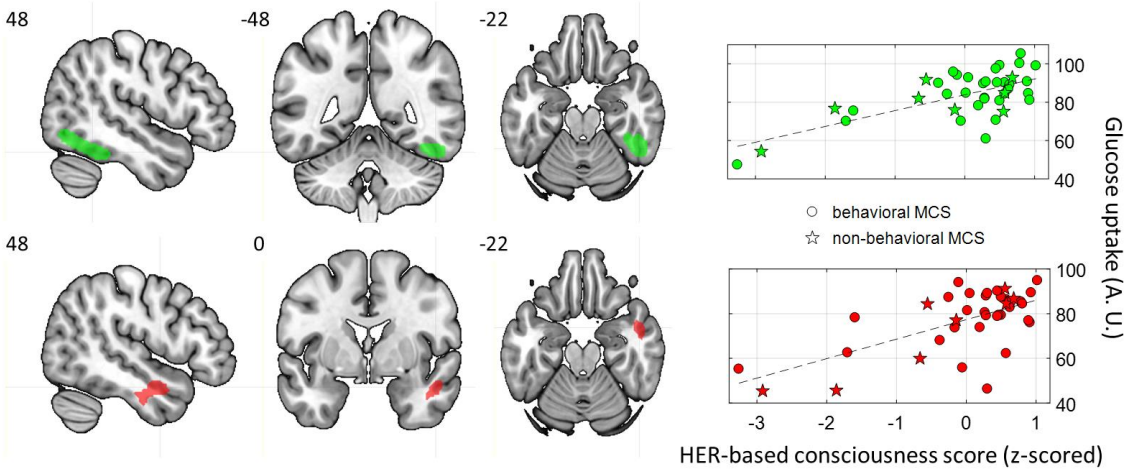


**Figure 3. Glucose metabolism correlates of HER-based consciousness scores in MCS patients (n=39).** Top, right ventral occipito-temporal regions (MNI [48, -48, -22]; FWE corrected, p-value cluster = 0.002). Bottom, right anterior superior temporal sulcus (MNI [48, 0, -22], FWE corrected, p-value cluster = 0.012).

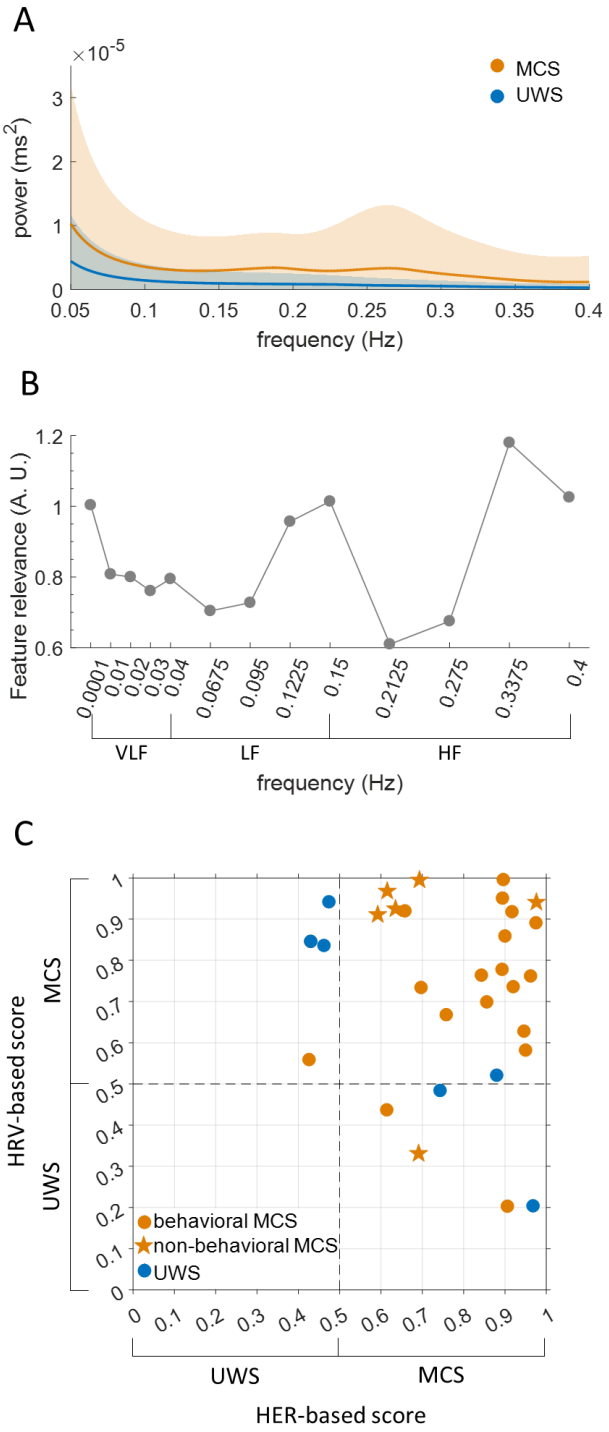


**Figure 4. Classifications based on the spectral analysis of heart-rate variability.**

(A) Heart-rate variability (HRV) power spectrum density of MCS and UWS patients in the training sample. (B) Relevance of each frequency (Gini impurity index) to the HRV-based classifier in the training sample, with most relevant frequencies at the border between low frequency (LF) HRV and high frequency (HF) HRV and at the upper limit of HF HRV. The

contribution of very low frequencies (VLF) appears limited. (C) Consciousness scores obtained from the classifier trained on HRV plotted against consciousness scores obtained from the classifier trained on heartbeat-evoked responses (HER), in the independent validation sample (n=30).

# Tables

**Table 1. Demographic information of patients included in the analysis.** TBI, traumatic brain injury.

| Group | PET Diagnosis | Number of patients (Females) | Mean age in years (range) | Mean months since onset. (range) | Etiology |
|---|---|---|---|---|---|
| Training | MCS | 31 (13) | 41 (18-73) | 33 (1-157) | TBI = 17, anoxia = 8, mix = 1, hemorrhage = 5 |
| | UWS | 7 (2) | 44 (28-65) | 23 (3-66) | TBI = 1, anoxia = 4, mix = 1, hemorrhage = 1 |
| Validation | MCS | 24 (13) | 38 (22-70) | 25 (2-118) | TBI = 9, anoxia = 9, mix = 1, hemorrhage = 2, infection = 2, hypoglycemia = 1 |
| | UWS | 6 (2) | 47 (32-65) | 35 (1-168) | Anoxia = 5, hemorrhage = 1 |
| Total | | 68 (30) | 38 (18-73) | 30 (1-168) | |

**Table 2. Summary of main results.** Classifications based on HERs consistently outperform classifications based on random EEG segments. Classifications using PET-based diagnosis consistently outperform classifications using CRS-R-based diagnosis. $p_{bin}$: p-value assessed from a cumulative binomial distribution ; $p_{mc}$: p-value from a Monte-Carlo permutation test.

| | | 3-fold cross validation accuracy, training sample | Monte-Carlo comparison between HER and random EEG segments | Accuracy in validation sample (test against chance, bin: binomial; mc: Monte-Carlo) | Monte-Carlo comparison between HER and random EEG segments |
|---|---|---|---|---|---|
| | HER 200-400ms | 81.62% | | 86.67% | |

| | | | | | |
|---|---|---|---|---|---|
| PET-based consciousness diagnosis | | | ] $p_{mc}$<0.001 | ($p_{bin}$<0.001, $p_{mc}$=0.001) | ] $p_{mc}$=0.064 |
| | Random EEG segments (mean ± SD across 1000 permutations) | 79.08%±0.7 | | 80.84%±4.09 ($p_{bin}$<0.001) | |
| CRS-R based consciousness diagnosis | HER 156-356 ms | 79.06% | ] $p_{mc}$=0.038 | 76.67% ($p_{bin}$<0.001, $p_{mc}$=0.004) | ] $p_{mc}$=0.041 |
| | Random EEG segments (mean ± SD across 1000 permutations) | 72.95%±4.49 | | 66.42%±7.13 ($p_{bin}$=0.272) | |

**Table 3. Significant brain regions that correlate with consciousness score in MCS patients.** Anatomical labels according to AAL Atlas (Tzourio-Mazoyer et al., 2002).

| Cluster | Brain regions | Number of voxels | Area ($mm^3$) | t-value peak | X | Y | Z |
|---|---|---|---|---|---|---|---|
| 1 | Temporal Inf R | 338 | 2704 | 5.494091 | 48 | -48 | -22 |
| | Fusiform R | 165 | 1320 | 5.415044 | 46 | -50 | -20 |
| | Occipital Inf R | 149 | 1192 | 5.088109 | 48 | -62 | -14 |
| | Lingual R | 13 | 104 | 4.532592 | 24 | -90 | -10 |
| 2 | Temporal Mid R | 86 | 688 | 4.845305 | 48 | 0 | -22 |
| | Temporal Inf R | 44 | 352 | 4.775336 | 44 | -10 | -36 |
| | Fusiform R | 18 | 144 | 4.723663 | 40 | -8 | -36 |
| | Temporal Pole Mid R | 3 | 24 | 4.513527 | 46 | 6 | -24 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Temporal Pole Sup R | 1 | 8 | 4.533496 | 46 | 2 | -20 |
| 3 | Temporal Sup R | 5 | 40 | 4.482989 | 64 | -22 | 0 |

**Table 4. Heart-rate variability univariate analysis**

| Heart-rate variability feature | MCS (mean ± SD) | UWS (mean ± SD) | Mann-Whitney U test p-value uncorrected for multiple comparisons | Rank sum | z-value |
|---|---|---|---|---|---|
| Mean interbeat intervals (ms) | 892 ± 194 | 817 ± 134 | 0.3661 | 629 | 0.9037 |
| Standard deviation of interbeat intervals (ms) | 47 ± 24 | 34 ± 23 | 0.1525 | 643 | 1.4309 |
| Root Mean Square of successive differences (ms) | 37 ± 29 | 19 ± 13 | 0.0768 | 652 | 1.7698 |
| Very low frequency power ($ms^2$) | 1.64 ± 4.04 $10^7$ | 3.83 ± 4.88 $10^7$ | 0.4743 | 585 | -0.7155 |
| Low frequency power ($ms^2$) | 468 ± 571 | 207 ± 276 | 0.1752 | 641 | 1.3556 |
| High frequency power ($ms^2$) | 637 ± 840 | 180 ± 189 | 0.1320 | 645 | 1.5062 |
| Low/High frequency ratio | 1.68 ± 2.09 | 1.54 ± 2.11 | 0.8213 | 611 | 0.2259 |