

# Supplementary Material: Cingulate and Cerebellar Beta Oscillations are Engaged in the Acquisition of Auditory-Motor Sequences

## Table of Contents for Supplementary Material

1. Supplementary Methods
2. Supplementary Results
3. Supplementary Figures

Figure S1. Page 10.

Figure S2. Page 12.

Figure S3. Page 13.

Figure S4. Page 14.

Figure S5. Page 15.

## 1. Supplementary Methods

### *Statistical Analysis*

In the statistical analysis of behavioural results using non-parametric paired permutation tests, we also report a non-parametric effect size estimator,  $PS_{dep}$ , following Grissom and Kim (2012).  $PS_{dep}$  is the probability that in a randomly sampled pair of *dependent* values (one matched pair: two values from the same participant under different conditions) the value from Condition B (which for instance has larger values) will be greater than the value from Condition A. The maximum value is  $PS_{dep} = 1$ . Since a paired permutation test permutes the sign of pairs of samples (sign test), we can proceed as follows: for two samples of length  $N$ , we first compute the difference between each of the  $N$  pairs of values from both samples, then we count the number of positive difference scores

$N_+$ . The probability of greater values in sample B relative to A is  $PS_{dep} = N_+ / N$ . If there are ties (zero difference), we reduce the denominator  $N$  by the number of ties  $N_0$  ( $PS_{dep} = N_+ / [N - N_0]$ ). A non-parametric estimation of effect size like  $PS_{dep}$  is more adequate when using non-parametric tests than reporting parametric effect size estimates such as Cohen's  $d$ , particularly because parametric effect size estimates are affected by deviations from normality and heterogeneity of variances.

### *Source reconstruction*

The CSP algorithm (Blankertz et al. 2008) is a method used to analyze multi-channel data based on recordings from two conditions. CSP leads to the generalized eigenvalue decomposition of the original signal  $x(t) \in \mathbb{R}^C$  into  $x_{CSP}(t) \in \mathbb{R}^C$ . The decomposition is parameterized by a matrix  $W \in \mathbb{R}^{C \times C}$  ( $C$  being the number of channels) as follows:

$$x_{CSP}(t) = W^T x(t) \quad (1)$$

Following e.g. Blankertz et al. (2008), we call each column vector  $w_j(t) \in \mathbb{R}^C$  ( $j = 1, \dots, C$ ) of  $W$  a *spatial filter* and each column vector  $a_j(t) \in \mathbb{R}^C$  of the inverse matrix  $A = W^{-1}$  a *spatial pattern*. The way in which matrix  $W$  is obtained follows an optimization criterion, such that CSP filters maximize the variance of the spatially filtered signal for one condition while minimizing it for the other condition.

The CSP components can be obtained by projecting the original sensor-space signals  $X$  using spatial filters in  $W$ :

$$Z = WX$$

The projected data in  $Z$  contain components (sources) and are sorted by the size of their

eigenvalue (from high to low).

Thus, the spatial patterns  $a_j$  can be viewed as a correlation map between original sensor signal  $x_i$  and the spatially filtered signal  $z_j$  (sources).

As in Nierula et al. (2013), we performed source reconstruction with current equivalent dipole modeling using the CSP patterns  $a_j$  associated with the three largest eigenvalues in each subject.

Statistics of inverse calculations: We tested the null hypothesis that across subjects the fitted dipoles could be located across all grid positions (in MNI space) with the same probability. To this aim we assumed that the probability of fitting a dipole to any point  $j$  within the individual warped grid was uniformly distributed ( $p_j = 1/N_{\text{grid}}$ , with  $N_{\text{grid}} = 37163$  the total number of grid points inside the subject-specific space). Next, we assessed the probability of grid points,  $p_{\text{loc}}$ , falling within each anatomical location from the AAL atlas. Our locations of interest were the cingulate gyrus, the temporal gyrus, the cerebellum, the SMA and the functional area of the dorsolateral prefrontal cortex (DLPFC): (contributing to the anatomical area of the middle frontal gyrus: labels Frontal\_Sup and Frontal\_Mid in the AAL atlas; Brodmann areas 9 and 46). For each of those locations, we treated the results as a *binomial experiment* consisting of a fixed number  $n$  of statistically independent Bernoulli trials ( $n = 20$  subjects), each with a probability of success  $p_{\text{loc}}$ , and we counted the number of successes  $k$  (meaning  $k$  subjects exhibited that location after dipole fitting). In a binomial experiment, the probability of  $k$  successes out of a sample of  $n$  independent variables (subjects) is determined by the binomial distribution

$$P(k,n) = \binom{n}{p_{\text{loc}}} p_{\text{loc}}^k q_{\text{loc}}^{(n-k)} \quad (2)$$

With  $q_{\text{loc}}$  being the probability of failure in the experiment  $q_{\text{loc}} = 1 - p_{\text{loc}}$ . Now, considering that for

each subject we fitted a dipole to each of the 3 CSP and assuming that these CSP are independent among them (implying that each CSP leads to a different location), the probability of success of one location being found for at least one of the 3 CSP is  $p_3 = 1 - (q_{loc})^3$  and  $q_3 = 1 - p_3$ . Accordingly, the probability that at least  $k$  subjects out of  $n$  have a source in the same specific location  $loc$  is

$$P(k,n) = \sum_{k'=k}^n \binom{n}{k'} p_3^{k'(n-k')} \quad (3)$$

The final p-value was corrected for multiple comparisons arising from the five locations of interest by using the Bonferroni correction:  $\alpha / 5 = 0.01$ , with  $\alpha = 0.05$ . Accordingly, the null hypothesis was rejected for any of the locations being tested whenever there was at least the minimum number of subjects  $k$  showing that same specific location, with a probability  $P(k,n)$  below the corrected significance threshold 0.01.

## 2. Supplementary Results

### *Questionnaire data*

After completing all performance blocks, participants filled out questionnaires asking them about their subjective error number estimates and their awareness of the different kinds of alterations of auditory feedback (AAF). The questions included in the questionnaire were:

Please, answer whether the following statements are CORRECT or WRONG.

1a. I produced fewer than ten errors per block by pressing a wrong key.

1b. I produced more than ten errors but fewer than twenty errors per block by pressing a wrong

key.

- 1c. I produced more than twenty errors per block by pressing a wrong key.
  
- 2. I realised that I made errors because the note was different than expected.
  
- 3. I realised that I made errors because the movement felt different than expected.
  
- 4. I memorized the sequence of keystrokes ignoring the tones.
  
- 5. I realised that the auditory feedback was occasionally modified.
  
- 6. I think that I always managed to realise when auditory feedback was externally modified.
  
- 7. I realised that some altered tones produced by the computer corresponded to elements from the sequence I was playing.
  
- 8. I realised that some altered tones produced by the computer were completely unrelated to the sequence I was playing.

Based on the answers provided in questions 1a-1c, we scored their estimated error number (per sequence type) as:

If 1a TRUE → 5 errors per sequence type

If 1b TRUE → 15 errors per sequence type

If 1c TRUE → 25 errors per sequence type

Participant #6 marked two options as true: 1b and 1c. Accordingly, we assigned an error number estimate of 20 to this participant.

The subjective estimate of self-produced error per sequence type was on average 19 (1), whereas the number of self-produced errors was 15 (2), and the difference was significant ( $p = 0.016$ ). Thus, participants overestimated the number of pitch errors due to the presence of AAF, an outcome that

has been reported in a previous behavioral study (Pfordrescher and Beasley, 2014).

The following table indicates the number of participants (N = 20) marking the questionnaire items

2-8 as true statements:

Questionnaire item	Number of participants selecting TRUE (Ntotal = 20)
2	9
3	18
4	14
5	18
6	12
7	13
8	16

Based on participant's responses, it seems that they might have preferred movement-based over auditory-based error detection (item 3 vs 2), as well as have aimed at learning the sequences without paying much attention to the auditory information. Significantly, however, the large majority of the participants retrospectively reported to have noticed the different kinds of AAF manipulations and have distinguished AAF from self-produced errors.

#### *Improvements in performance during the familiarization session*

During the familiarisation stage, participants exhibited significant improvements in performance timing across training trials (1 to 3): reduced average tempo, 420 [10] ms and 385 [10] ms,  $p = 10^{-3}$ ,  $PS_{dep} = 0.95$ ; reduced temporal variability 0.28 [0.03] to 0.25 [0.03],  $p = 0.01$ ,  $PS_{dep} = 0.75$ .

In addition, in the first trial of the performance block (normal feedback) there were additional improvements in performance relative to the last trial of the training block: reduced average tempo, 384 [9] ms and 378 [9] ms,  $p = 0.03$ ,  $PS_{dep} = 0.70$ ; reduced temporal variability, 0.25 [0.03] to 0.23

[0.02],  $p = 0.04$ ,  $PS_{dep} = 0.65$ .

Thus, there were initial improvements in performance during the familiarisation phase, and further improvements in the first normal trial of the performance session. However, the introduction of AAF during the subsequent trials of this session disturbed learning. This converges with previous evidence supporting that a distractor task can suppress the performance changes associated with learning (Seidler et al., 2002).

### *Pitch errors induced by AAF*

We assessed the properties of the self-produced pitch errors which may have been induced by AAF events (here termed 'AAF-induced-errors'). To this aim, we selected pitch errors that followed an AAF event and preceded the next one, with the constraint of a maximum distance of 10 keystrokes from previous AAF. Note that the average rate of AAF was one in every 8.37(0.05) keystrokes (same for ASO and UAF trials). Our stimulus material were sequences of 4 or 5 elements. Specifically, we looked into the three following properties:

1. Distance in number of keystrokes between the error and the previous AAF .

The AAF-induced-errors occurred between  $n+3$  and  $n+5$  keystrokes from the AAF (on average at 4.3 [0.1] subsequent keystrokes after ASO, and 4.5 [0.1] after UAF; n.s. difference between position in ASO and UAF trials,  $p > 0.05$ ). This outcome indicated that AAF induced pitch errors at the same ordinal position of the current AAF (or one position before, i.e. anticipating a potential upcoming AAF) but in the next rendition of the sequence.

2. Distance in ordinal position between the event in which an AAF was introduced and the subsequent AAF-induced-error (Figure S4A).

The majority of the AAF-induced-errors, which occurred in the next rendition of the sequence, fell at the same (distance = 0) or at one earlier ordinal position (distance = -1) than the previous AAF event. In the latter case, the self-produced error anticipated with a lag 1 a potential repetition of an AAF event at the same ordinal position as it fell in the previous rendition. There was no significant difference between the proportion of distance values in ASO and UAF trials ( $p > 0.05$  for all values).

3. Distance between the ordinal position that the AAF-induced-error activated (i.e. the wrong pitch activated an action of the sequence corresponding to an ordinal position 'ordpos\_error') and the ordinal position in which the error occurred (i.e. the event which should have been played: 'ordpos\_correct', See Figure S4B).

The largest proportion of the variable 'ordpos\_error - ordpos\_correct' was for a difference of 1, and the proportion was significantly larger in ASO relative to UAF trials (0.48 [0.03] and 0.40 [0.02], respectively,  $p = 0.021$ ). This outcome indicated that the majority of the AAF-induced-errors were due to the anticipation of the pitch value of the next sequence element.

In sum, our detailed analysis of the properties of the AAF-induced-errors revealed the following main outcomes:

Participants compensated for the previous presence of (unpredicted) altered feedback by modifying the sequence contents in the next rendition. Specifically, the compensation mostly consisted of replacing the sequence element at the same or prior ordinal position in which the previous AAF fell ( $m = n, n-1$ ; Figure S4A) with the anticipated subsequent sequence element ( $m = n+1$  or  $n$ , respectively; Figure S4B). Moreover, this process occurred significantly more often following ASO events. Taken together these results indicate that participants altered the sequential organization of



the events planned for production to compensate for the previous AAF, and most often in the case of altered feedback simulating serial-order errors (ASO events). These outcomes are consistent with previous behavioral studies on sequential planning (e.g. Pfordrescher and Palmer, 2006), yet they are novel in demonstrating that – at least for short sequences – compensatory changes primarily occur in the next rendition of the sequence.

#### *Additional results of the source localisation analysis*

Here we report additional sources that were localised in a smaller number of participants following our CSP + dipole fitting procedure. These additional sources, however, were non-significant according to our statistical analysis at the group level. Below we present sources that were found in at least 5 participants.

For the ASO – NAF feedback comparison, additional sources of enhanced theta-band activity were located in a few subjects to the precentral gyrus (contralateral primary motor cortex, 8/20 subjects), dorsolateral and medial prefrontal cortex (7/20 and 6/20 subjects respectively), postcentral gyrus (contralateral somatosensory cortex, 6/20 subjects) and the thalamus (8/20 subjects).

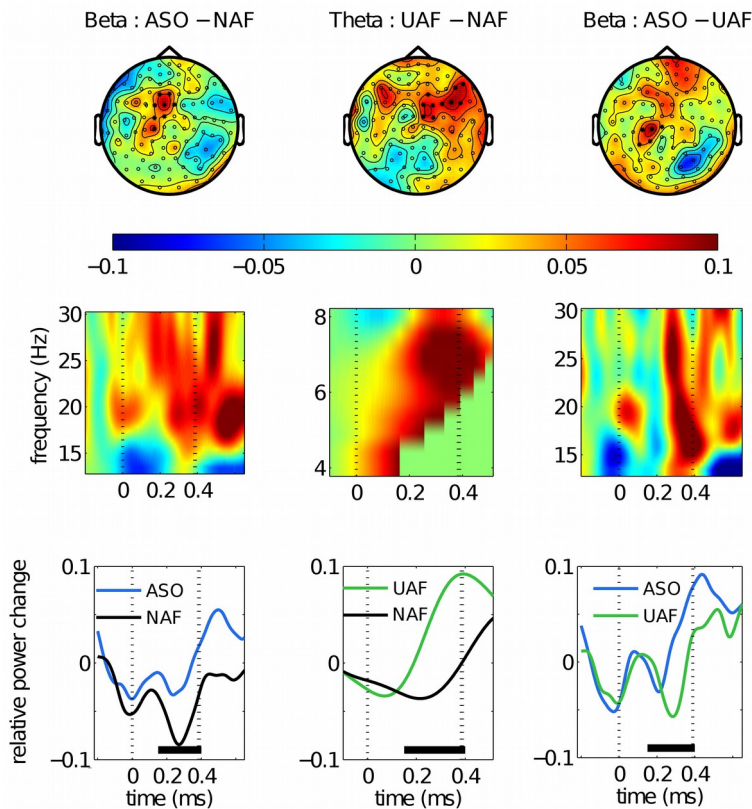
For the UAF – NAF feedback comparison, additional sources of enhanced theta-band activity were located in a few subjects to the dorsolateral and medial prefrontal cortex (9/20 and 7/20 subjects respectively), the thalamus (7/20 subjects) and the cingulate cortex (6/20 subjects).

Finally, the effect of enhanced beta-band oscillatory activity obtained for the specific comparison between ASO and UAF trials was localized in a few subjects to the occipital lobe (9 / 20 subjects), the thalamus (8/20 subjects) and the temporal lobe (6/20 subjects).”

Finally, the effect of enhanced beta-band oscillatory activity obtained for the specific comparison

between ASO and UAF trials was localized in a few subjects to the occipital lobe (9 / 20 subjects, n.s.), the thalamus (8/20 subjects, n.s.) and the temporal lobe (6/20 subjects, n.s.)

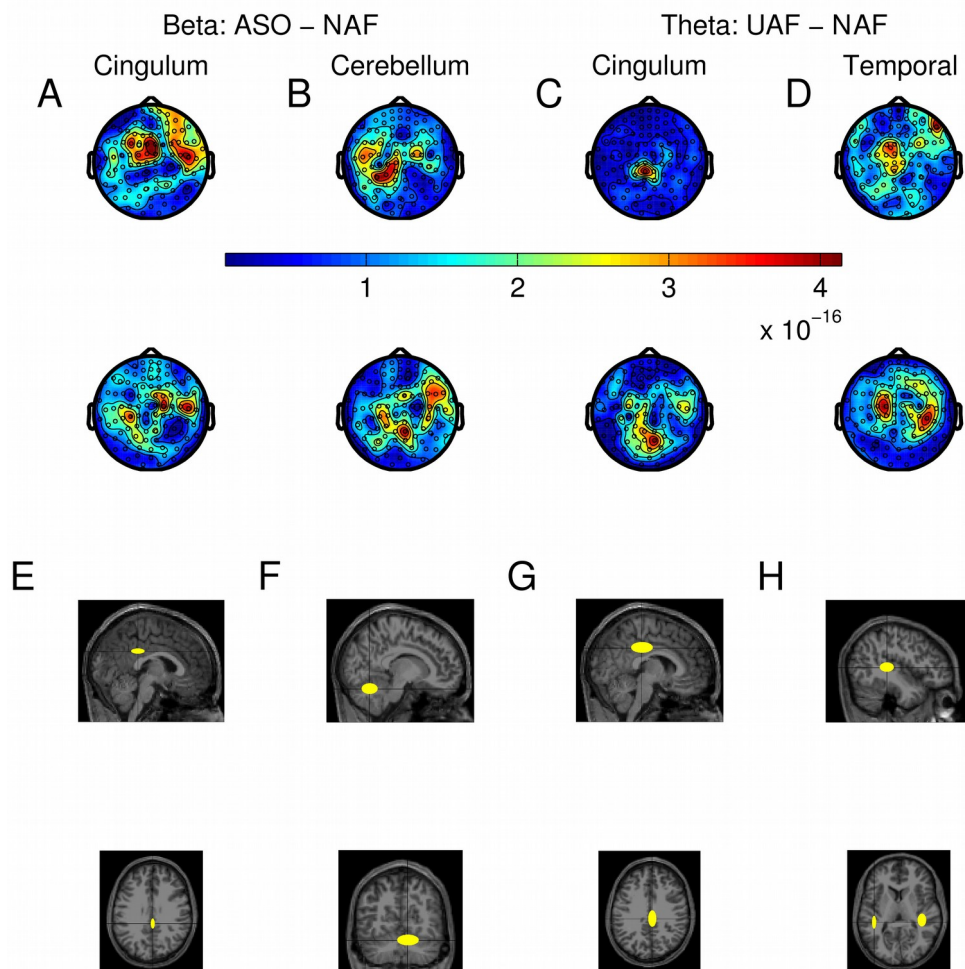
### 3. Supplementary Figures



**Figure S1. Effects of alterations of auditory feedback (AAF) on oscillatory power in the sensor-space.**

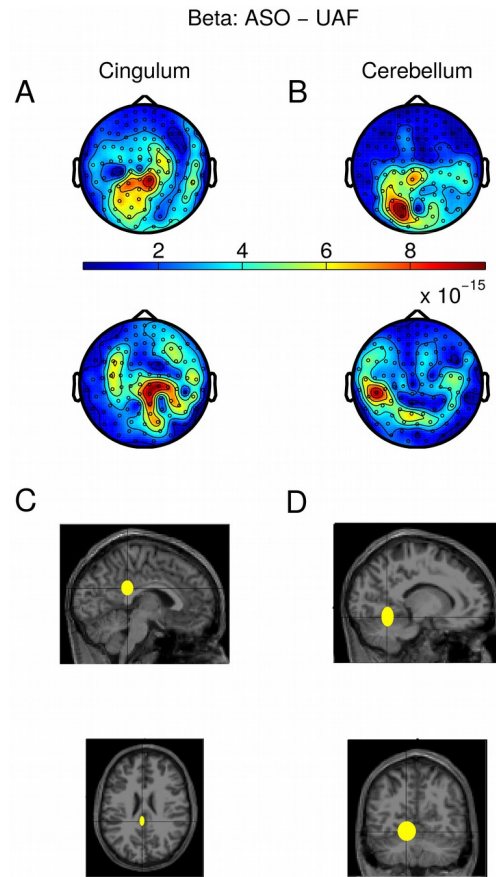
Planar gradiometers. Top row shows scalp topographies for relative power changes in the beta (left & right) and theta band (center), corresponding to the significant clusters obtained within 150 to 400 ms (0 ms is keystroke and auditory feedback onset; cluster permutation test,  $p < 0.025$ , two-sided test) for the different between-condition comparisons. Topographies are displayed in a combined planar gradiometer representation. The black stars denote the sensors belonging to the significant clusters. The left map presents the comparison between trials with serial order alterations (ASO) and normal feedback (NAF), revealing a significant positive cluster in the beta frequency region with a *midline frontocentral* scalp

distribution. The center map reveals a significant positive cluster in the theta band with a *right frontal* scalp distribution, corresponding to the comparison between trials with unrelated auditory feedback alterations (UAF) and NAF. The right map displays a comparison between ASO and UAF trials, demonstrating a significant positive cluster in the beta band, which had a *central* scalp distribution. Middle row shows grand-averages of the time-frequency power changes over the sensors pertaining to the significant positive clusters shown above. Lower row shows some time courses of the cluster-based power averaged within the corresponding significant frequency band as shown in the middle row.

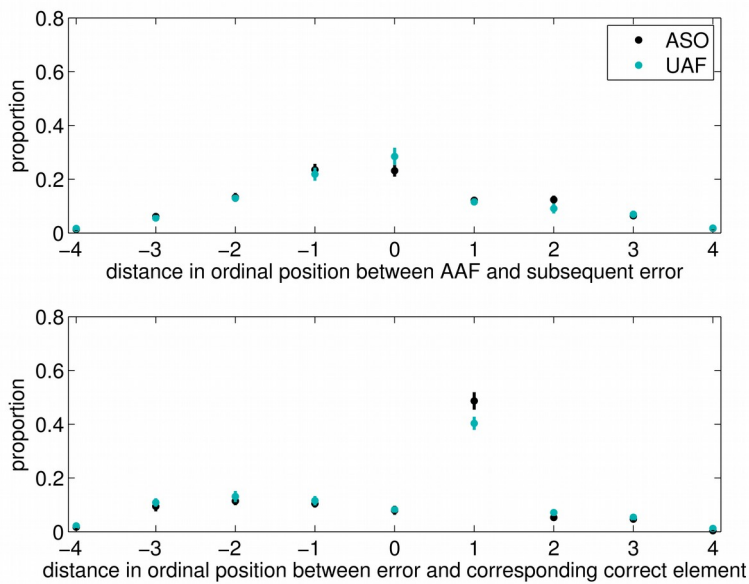


**Figure S2. CSP and source localization of beta and theta oscillatory activity in planar gradiometers for the of ASO-NAF and UAF-NAF difference, respectively. A-B. Beta-band CSP patterns (in arbitrary**

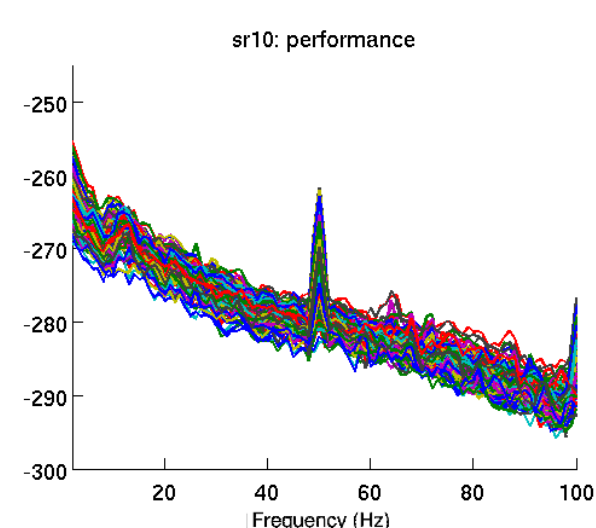
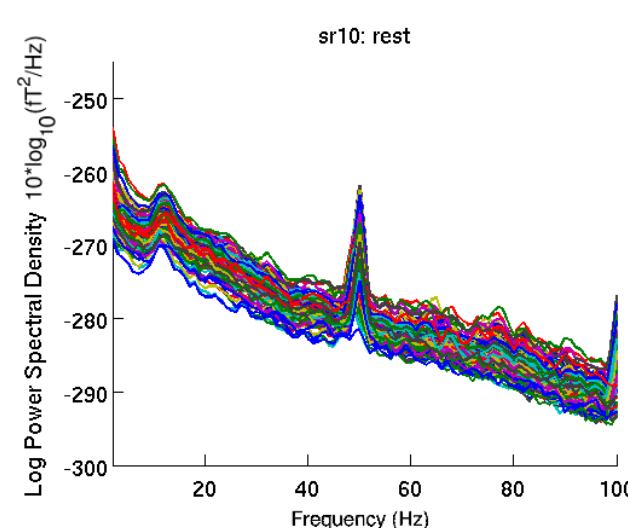
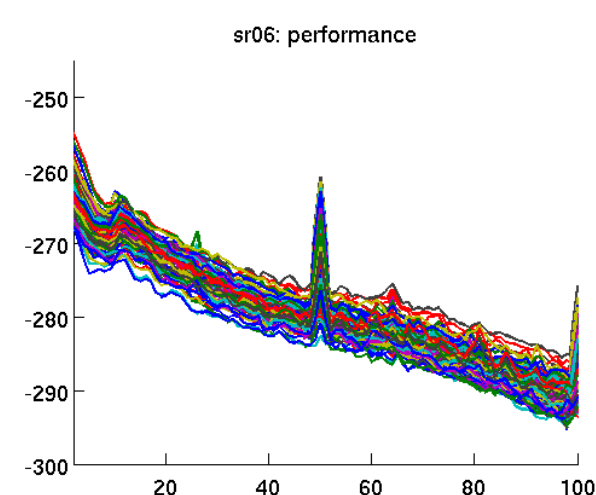
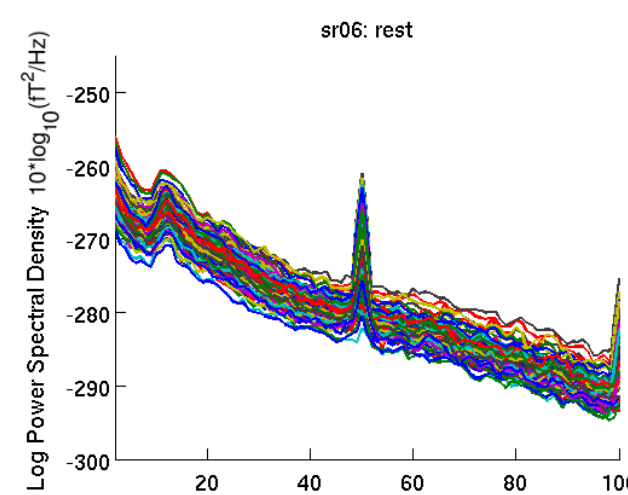
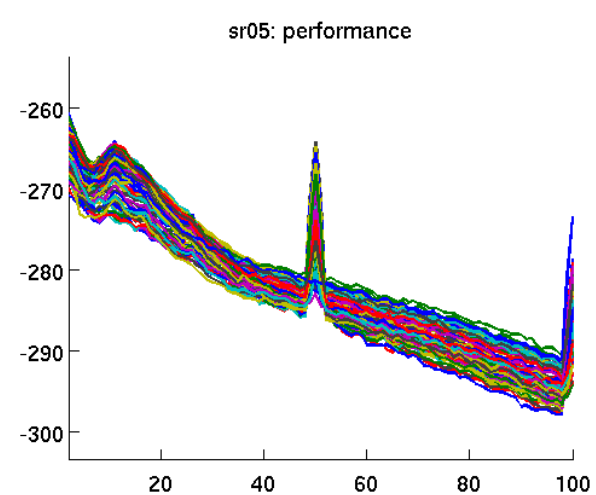
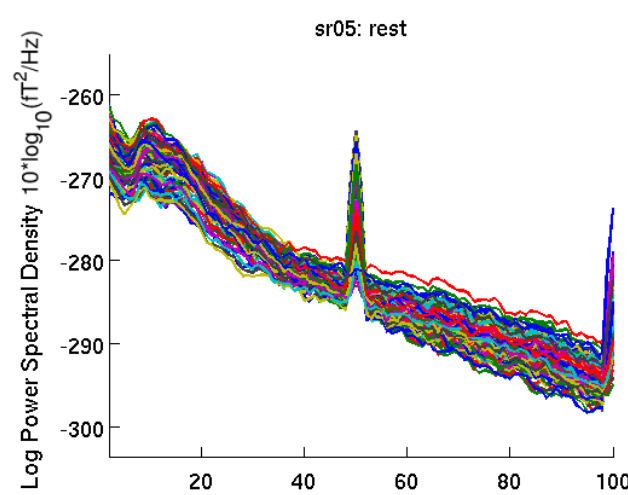
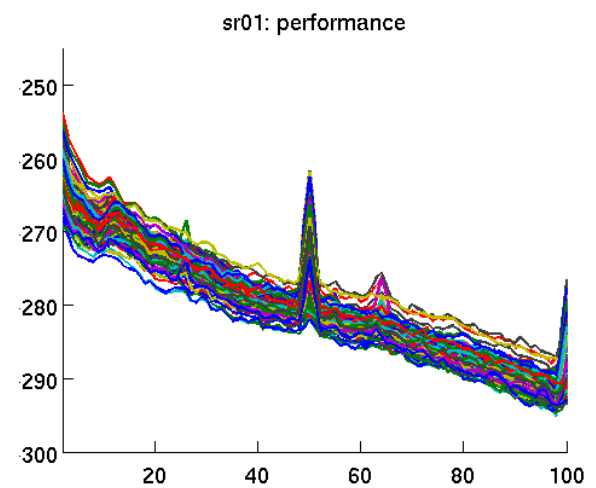
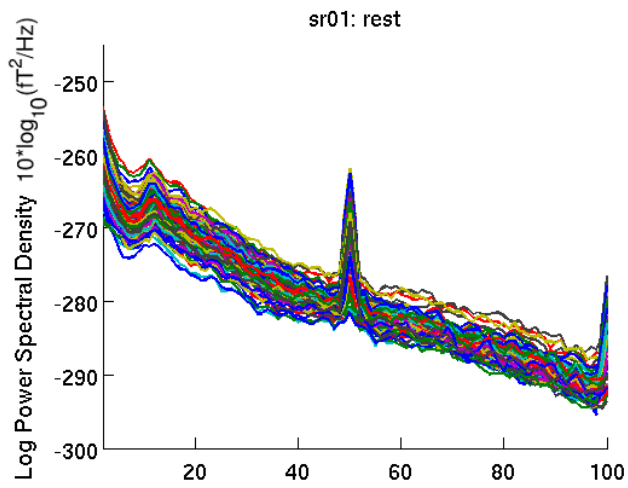
units, a.u.) obtained for one of the three largest eigenvalues from two representative subjects (upper row: subject #5, lower row: subject #10). These CSP patterns corresponded to an enhanced beta-band oscillatory activity in the ASO as compared to the normal feedback condition and were generated by a monopolar source in the cingulum (A) or cerebellum (B). The patterns are displayed in a combined planar gradiometer representation. C-D. Theta-band CSP patterns (in a.u.) obtained for one of the three largest eigenvalues from the same representative subjects as in (A-B). These CSP patterns corresponded to an enhanced theta-band oscillatory activity in the UAF as compared to the normal feedback condition and were generated by a monopolar source in the cingulum (A-B) or a dipolar source in the temporal gyrus (C-D). (E-H) Standard MNI coordinates (median and median absolute dispersion across subjects) of the anatomical locations of the dipoles generating the CSP patterns obtained for each between-condition difference in the sensor space (represented in the upper panels).



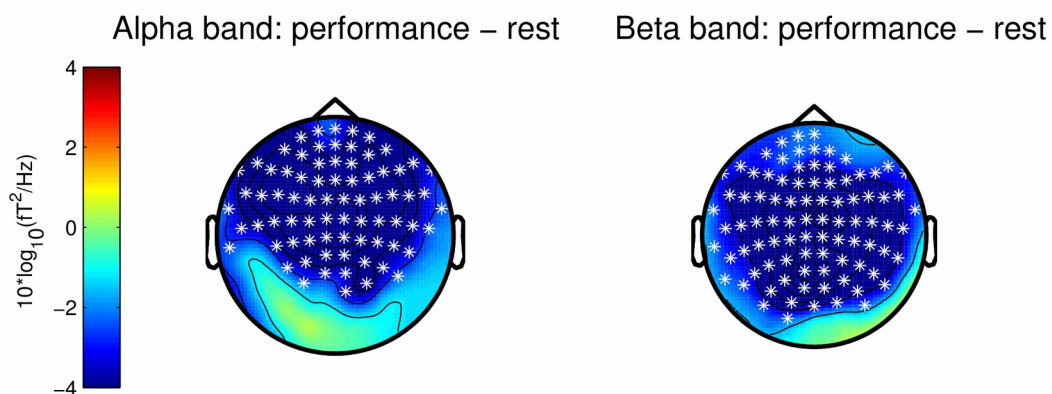
**Figure S3. CSP and source localization of beta oscillatory activity in planar gradiometers for the of ASO-UAF difference.** (A-B) CSP patterns (in a.u.) from two representative subjects (upper row: subject #12; lower row: subject #17) in association with one of the three largest eigenvalues, and corresponding to the maximization of variance of beta oscillatory activity in the modified feedback condition ASO as compared to UAF. The representative beta-band CSP patterns were generated by sources in the cingulate gyrus (A) and the cerebellum (B). The patterns are displayed in a combined planar gradiometer representation. (C-D) The anatomical locations of the dipoles generating the beta-band CSP patterns are provided as median and median absolute dispersion across subjects in standard MNI coordinates.



**Figure S4. Proportion of values (and SEM bars) for the following variables:** (A) Distance in ordinal position between the AAF event and the subsequent self-produced error (which occurred in the next rendition of the sequence); (B) Distance between the ordinal position that the wrong pitch of the AAF-induced-error activated ('ordpos\_error') and the ordinal position in which the error occurred ('ordpos\_correct'):  $\text{ordpos\_error} - \text{ordpos\_correct}$ .



**Figure S5. Power spectral density (PSD, in units of  $10 \cdot \log_{10} [fT^2/Hz]$ ).** The PSD of the magnetometers in the performance blocks was estimated separately for (i) the *resting* periods between performance trials (amounting to approximately 2 minutes per subject; left panels) and (ii) the *performance* trials (15 trials x 23 seconds amounts to 5.75 minutes; right panels). The four rows depict the PSD results for four representative participants with sources localised to the cerebellum. No increase in PSD in the higher frequency range above 20Hz is observed during performance, relative to rest recordings, supporting that there was not a significant contribution of muscle-artifacts to the general level of oscillatory activity. **See Figure S6 for statistical analyses of these effects.** The most notable difference between the PSD during rest and performance was a higher level of alpha (8-13Hz) power at rest, as expected.



**Figure S6. Significant clusters of differences between performance and rest in power spectral density (PSD) in the alpha (8-12Hz) and beta ranges (13-30Hz).** Scalp topographies for PSD changes (in units of  $10 \cdot \log_{10} [fT^2/Hz]$ ) in the alpha (left) and beta (right) frequency ranges, corresponding to the significant clusters obtained (cluster permutation test,  $p < 0.025$ , two-sided test) for the between-condition comparisons. In both cases a negative cluster was found due to the significantly lower PSD during performance relative to rest. **A similar analysis performed in the gamma range (31-100Hz) revealed only a trend for significance ( $p = 0.045$ , non-significant in our two-sided test) due to less pronounced gamma power during performance than during rest over parietal and occipital electrodes.**