Supplemental Materials

Journal section: Behavioral / Systems / Cognitive

Decoding and cortical source localization for intended movement direction with MEG

Abbreviated title: Decoding and source localization with MEG

W Wang[1, 2, 3, 4], G P Sudre[2, 3, 5], Y Xu[3,9], R E Kass[3, 6], J L Collinger[1], A D Degenhart[2, 3], A I Bagic[7, 8], and D J Weber[1, 2, 3, 4,]

[1]Department of Physical Medicine and Rehabilitation, University of Pittsburgh, Pittsburgh, PA, USA
[2]Department of Bioengineering, University of Pittsburgh, PA, USA
[3]Center for the Neural Basis of Cognition, Pittsburgh, PA, USA
[4]Quality of Life Technology (QoLT) Engineering Research Center, Pittsburgh, PA, USA
[5]Program in Neural Computation (PNC), Carnegie Mellon University, Pittsburgh, PA, USA
[6]Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, USA
[7]Department of Neurology, University of Pittsburgh, PA, USA
[8]Center for Advanced Brain Magnetic Source Imaging (CABMSI), Pittsburgh, PA, USA
[9]Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA, USA

Corresponding author: Douglas J Weber
3471 Fifth Ave, Suite 202
Pittsburgh, PA 15213
djw50@pitt.edu

Number of supplemental figures: 4
Number of supplemental tables: 1

Linear Discriminant Analysis

The linear discriminant analysis finds a linear projection of data which maximizes the ratio of between-class scatter and within-class scatter. For a multi-class problem, the between-class and within-class scatter matrices are

$$S_B = \sum_C N_C (\mu_C - \mu)(\mu_C - \mu)^T$$

$$S_W = \sum_C \sum_{i \in C} (x_i - \mu_C)(x_i - \mu_C)^T$$

where

$$\mu_C = \frac{1}{N_C} \sum_{i \in N_C} x_i$$

$$\mu = \frac{1}{N} \sum_{i \in N} x_i$$

and where $c \in [1, 2, 3, 4]$ indicates 1 of 4 possible movement directions; N_c is the number of trials in direction c; N is the total number of trials; x_i is a data vector of length m (number of motor sensors: 87), where each entry is the recorded MEG signal averaged over the decoding window for one sensor. Then, LDA finds a projection matrix $\mathbf{W}^*$ such that

$$W^* \leftarrow \arg\max_W \frac{\left|W^T S_B W\right|}{\left|W^T S_W W\right|}$$

It can be shown that the solution of $\mathbf{W}^*$ is a m × k matrix where each column $\mathbf{w}_i$ is an eigenvector from the eigenvalue equation

$$S_W^{-1} S_B w_i = \lambda_i w_i$$

where $\{\lambda_i \mid i = 1, 2, ..., k\}$ corresponds to k eigenvalues. However, there are maximal c − 1 non-zero eigenvalues and hence the highest rank of the projection matrix is c − 1 (i.e. 3 in this case since there are 4 directions). To obtain the projected low-dimensional data Y from original data matrix X, later used for decoding, the following transformation was used:

$$Y = XW^*$$

Note that the matrix $\mathbf{S}_w$ only becomes singular when the dimension of the data exceeds the number of trials, which was never the case for the current analysis since there were always more trials than dimensions.

Discussion on the number of bootstrap draws

The bootstrap procedure we have used here is somewhat unusual. The usual bootstrap ANOVA tests the null hypothesis that the observations follow the same distribution across conditions, and it does this by shuffling the observations. Here the "observations" that are shuffled are the sensor signals and, instead, we test the null hypothesis that the source signals follow the same distribution across conditions. Each source signal (an image) is a statistic, meaning a function of the observations. In theory, if we knew the null distribution of each source pixel image and could generate observations from it, we would test the null hypothesis by generating G observations for each condition and running an ANOVA. If the null distribution were exactly normal, then it would not matter how big G would be, because the ANOVA F distribution would be exact.

There are two parts to the bootstrap test we have applied. First, we use the standard bootstrap argument to say that, for each condition, as the number of trials increases the empirical cumulative distribution function (CDF) will converge to the theoretical CDF; and, for each condition, with large numbers of trials the bootstrap distribution of the images will converge to the theoretical distribution of the images. Second, we use the standard propagation of uncertainty (delta method) argument to ensure that, for each condition, as the number of trials increases the source image will converge to a normal distribution. Combining these, for large numbers of trials the bootstrap distribution of the images will be approximately the same as the theoretical distribution and both will be approximately normal.

Therefore, regardless of the number of bootstrap samples, for sufficiently many trials, the ANOVA F distribution, and the resulting p-value, will be approximately correct. In practice, the bootstrap distribution and the theoretical distribution may deviate somewhat from normality. Because the ANOVA F distribution is generally fairly robust, some departure from normality is not worrisome as long as G is not very small. In any case, we have examined some of the bootstrap distributions empirically and found them to be very nearly normal. Therefore, the p-values should be very nearly correct with G=50 bootstrap draws per condition. In our case we used a likelihood ratio test instead of the standard ANOVA because the variances were clearly unequal, which would violate the usual ANOVA assumptions. However, the argument concerning the validity of the p-value remains the same.

Discussion on the likelihood ratio test and the MANOVA test

The likelihood ratio test follows Section 2.1 from Behseta et al. 2007. Here point-wise likelihood ratio test is performed over individual cortical sources where each source signal is averaged within a specific time window. The null hypothesis is that the mean signal (averaged over repeated trials) at a source is equivalent under all conditions (in this case across four movement directions). Then it can be shown that the resulting test statistic follows a chi-squared distribution with C-1 degrees of freedom where C is the number of conditions. Since this likelihood ratio test does not rely on the assumption of equal variances in cortical activity across different movement directions, we used this test to maximize our power to identify individual cortical sources that are modulated by movement direction. We also repeated the same procedure using the standard ANOVA and tested movement modulation of individual cortical sources. The result is similar to that obtained from the likelihood ratio test. In the subsequent source space data analysis, we used MANOVA to characterize the temporal dynamics of cortical representation of movement direction across all cortical sources. MANOVA is a well established method commonly used to examine whether multiple dependent variables (cortical source activities) differ for different task conditions (four movement directions).

Discussion on excluding subjects based on movement direction decoding using EOG

EOG signals produced during eye movements may introduce artifacts in the MEG, particularly in the MEG sensors located nearest to the eyes. For example, if a subject made eye movement toward the target (i.e. in the intended movement direction) and the eye movement contaminated MEG recordings, it is possible to observe MEG signals that seem to be modulated by "intended movement direction" and to even decode movement direction from MEG signals. One approach to determine if EOG has at least partially contributed to the capability of decoding movement direction from MEG signals is to examine whether intended movement direction can be decoded from EOG, i.e. whether EOG itself actually contains information about intended movement direction. If intended movement direction information cannot be decoded from EOG (decoding accuracy falls within the 95% confidence interval of chance level, which is 25% for four-direction classification), this indicates that EOG does not contain information about intended movement direction and that EOG's contribution to movement direction decoding from MEG signals is minimum. We performed decoding analysis using EOG signals instead of MEG signals within the same decoding time windows. For nine out of ten subjects, EOG signals didn't contain information about intended movement direction. For only one subject (Subject 10), intended movement direction can be decoded from EOG signals. Hence, we decided to exclude that subject from our data analysis to eliminate the possibility of using MEG data contaminated by EOG artifact.

Discussion on readiness field (RF) and readiness potential (RP)

The current study focuses on the time-domain signal of MEG recording, and an earlier study has suggested that the premovement time domain signal is related to the concepts of readiness field (RF) and readiness potential (RP) (Waldert et al. 2008). The RP is an electrically negative waveform preceding volitional movement, and it was first recorded with scalp electroencephalography (EEG) by Kornhuber and Deecke in 1965. Later, using magnetoencephalography (MEG), Deecke et al. (1982) discovered the RF. Their study observed similar time courses and morphologies between RF and RP. Subsequently, there are more studies on RF (e.g. Pederson et al. 1998; Takahashi et al. 2004); however, the exact linkage between RF and RP has not been elucidated. The RF and RP are likely produced by neuronal activities at different cortical structures in close vicinity, with the RF (the magnetic signal) from tangential dipoles located in the sulci and RP (the electrical signal) more dominantly from radial dipoles located in the gyri (Sharon, Hämäläinen, et al. 2007; Hämäläinen et al. 1993). Therefore, RF should probably not be viewed simply as the magnetic field induced by the same electrical current that gives rise to RP.

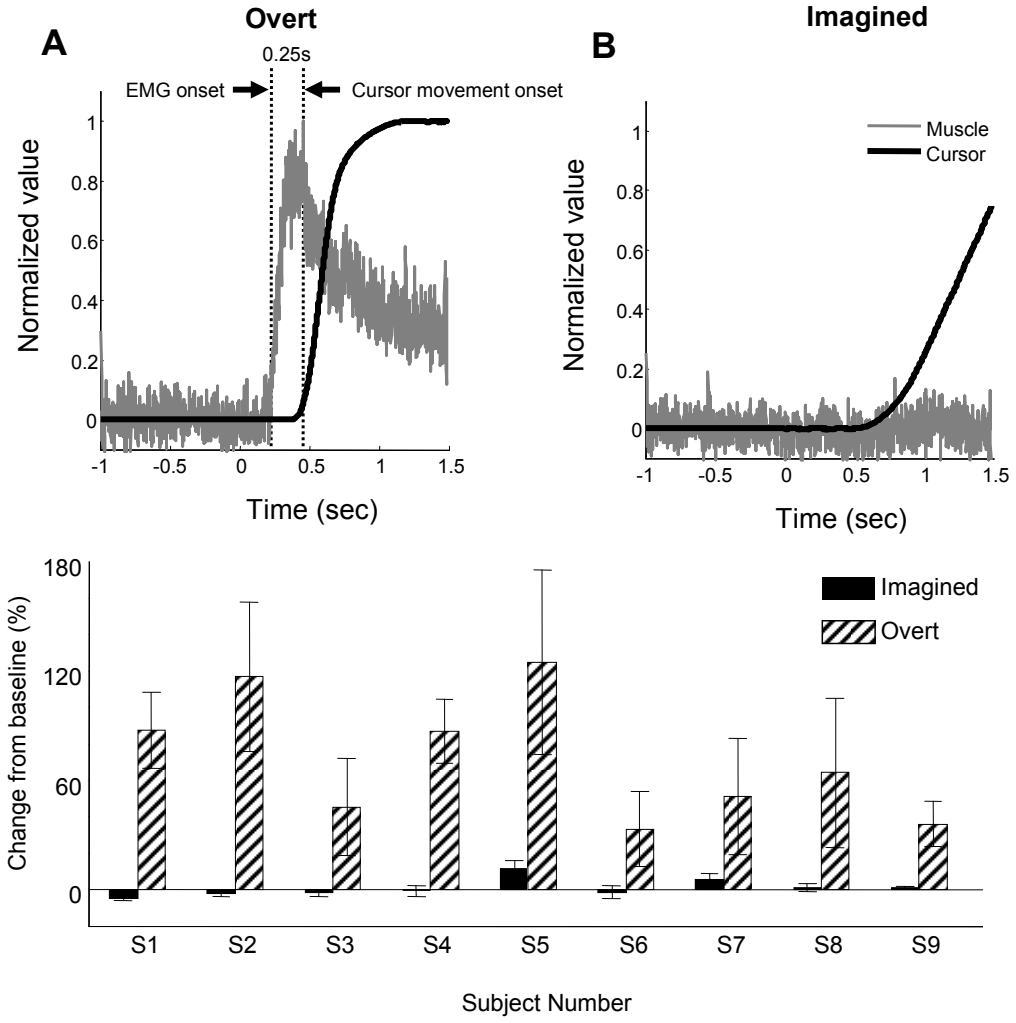
However, their close resemblance in timing and morphology may be a reflection of the activation of a common cortical area (e.g. the primary motor cortex) spanning both the gyri and sulci.

Related references:

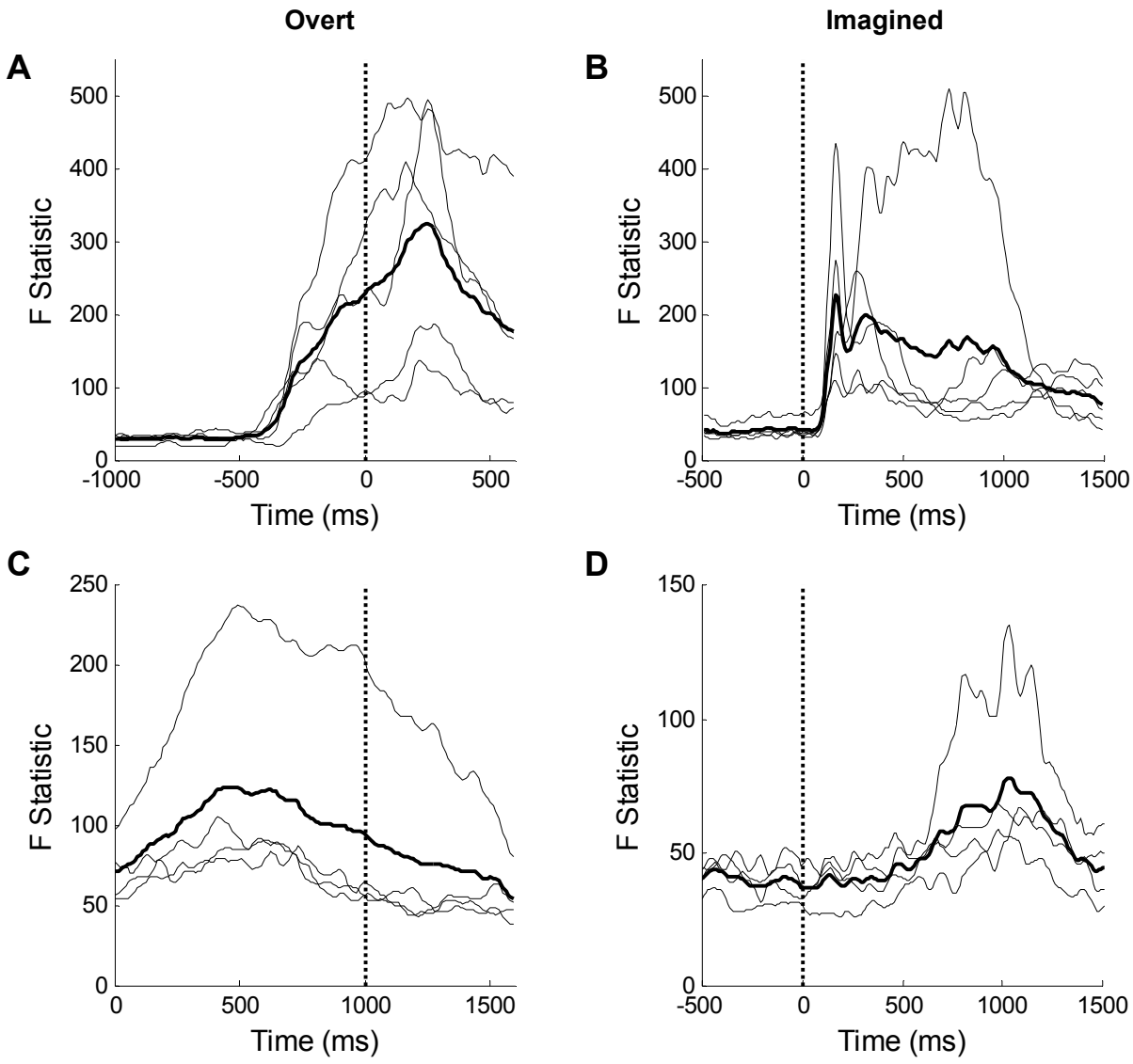
1. Deecke L, Weinberg H, and Brickett P. Magnetic fields of the human brain accompanying voluntary movement: Bereitschaftsmagnetfeld. Exp Brain Res 48: 144-148, 1982.
2. Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev Mod Phys 65:413– 497.
3. Kornhuber HH, and Deecke L. Changes in the Brain Potential in Voluntary Movements and Passive Movements in Man: Readiness Potential and Reafferent Potentials. Pflugers Arch Gesamte Physiol Menschen Tiere 284: 1-17, 1965.
4. Pedersen JR, Johannsen P, Bak CK, Kofoed B, Saermark K, and Gjedde A. Origin of human motor readiness field linked to left middle frontal gyrus by MEG and PET. Neuroimage 8: 214-220, 1998.
5. Sharon D, Hämäläinen MS, Tootell RB, Halgren E, and Belliveau JW. The advantage of combining MEG and EEG: comparison to fMRI in focally stimulated visual cortex. Neuroimage 36: 1225-1235, 2007.
6. Takahashi M, Watanabe Y, Haraguchi T, Kawai T, Yamane GY, Abe S, Sakiyama K, Hiraide Y, Lee WH, Ide Y, and Ishikawa T. Neuromagnetic analysis of the late phase of the readiness field for precise hand movements using magnetoencephalography. Bull Tokyo Dent Coll 45: 9-17, 2004.
7. Waldert S, Preissl H, Demandt E, Braun C, Birbaumer N, Aertsen A, and Mehring C. Hand movement direction decoded from MEG and EEG. J Neurosci 28: 1000-1008, 2008.

Supplemental Table 1. Decoding accuracy (%) relative to movement onset. Decoding results for target direction in the overt task for subjects without delay. The decoding window was chosen such that it always preceded movement onset by different time periods, which have been shown to be acceptable values for the delay between cortical activity and the start of movement (Moran and Schwartz, 1999). The length of the decoding window was 100ms for every subject based on the peak of the MANOVA curve (see Methods section). Chance accuracy is 25%. Accuracies are still above chance level even when decoding with data 250ms before movement onset. By doing that, the period containing EMG activity is not used for decoding, but it also ignores any motor cortical activity that occurs concurrently to that EMG activity, which decreases the results compared to what was shown in the paper. For ease of comparison, the third row shows decoding accuracies using decoding windows right before movement onset (copied from Table 1 overt movement condition in the main text).

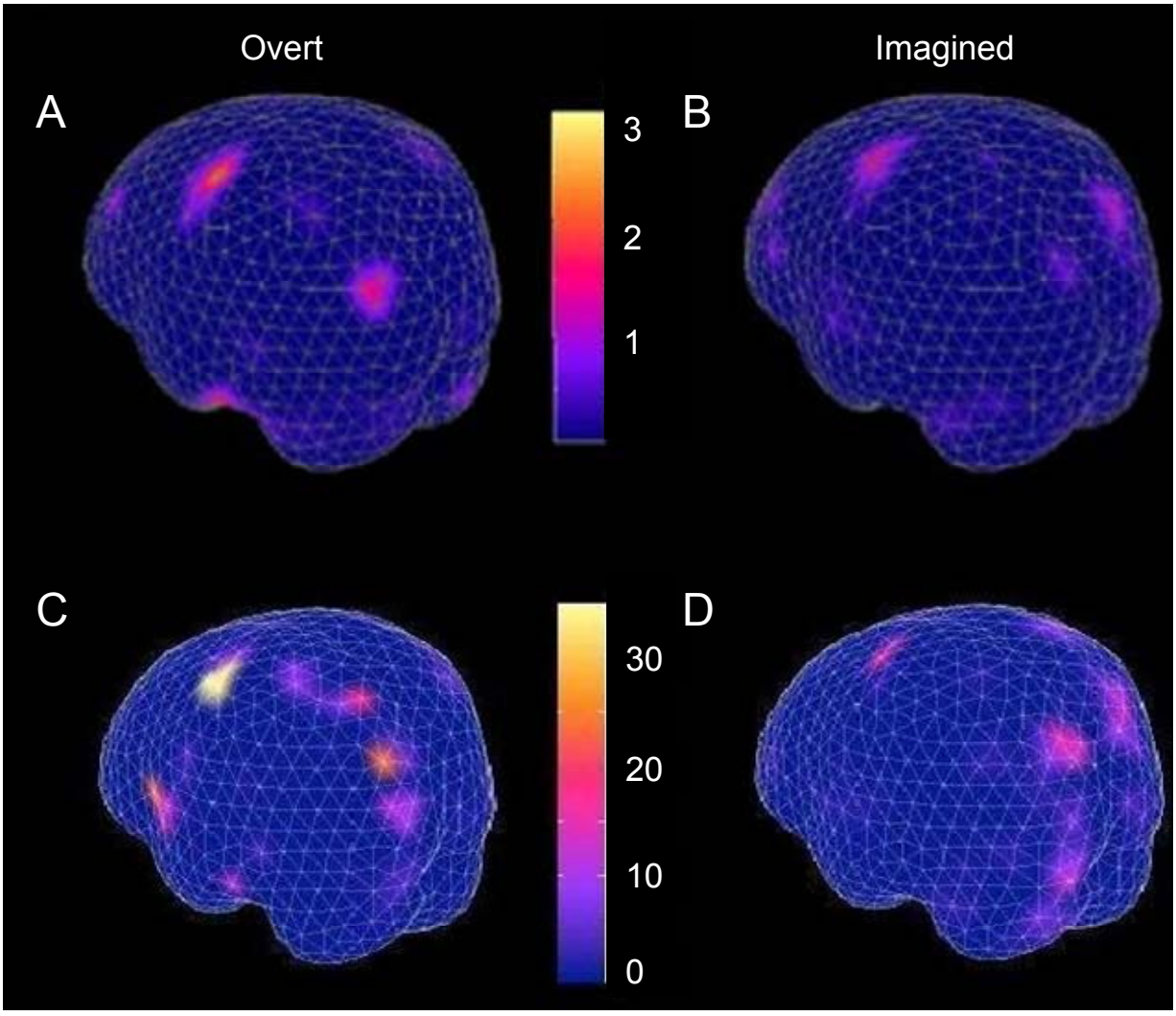
End of decoding window	S1	S2	S3	S4	S5	Mean
250 ms before movement onset	70.1	54.2	60.3	57.1	33.9	55.1
150 ms before movement onset	81.2	65.5	70.2	65.7	48.0	66.1
0 ms before movement onset	88.6	69.0	72.4	77.8	57.9	73.1



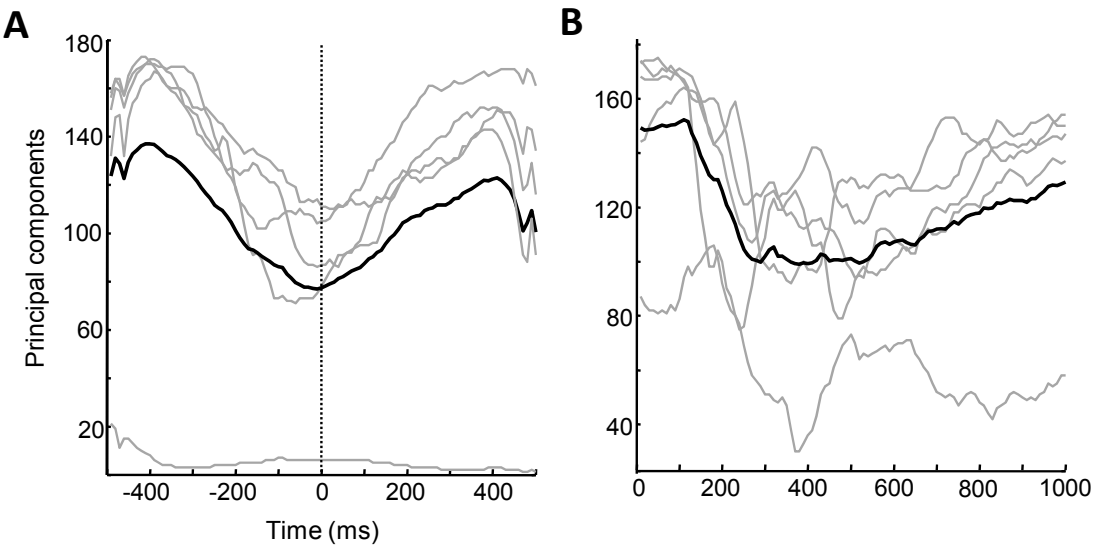
Supplemental Figure 1. EMG activity for both overt and imagined movements. EMG activity shown in this figure is the total muscle activity summed over wrist flexor and extensor muscles (flexor carpi radialis and extensor carpi radialis). *A & B*: These two plots show EMG activity and cursor movement for overt and imagined movement, respectively (Subject S4, rightward movement, Time 0 is target onset). EMG signals are rectified and averaged over all trials (blue curves), and the baseline EMG activity during rest is removed. Cursor movement (displacement from the center target) is also averaged over all trials (red curves). EMG and cursor movement for overt and imagined movement are normalized to their respective maximum value during overt movement. Onsets of EMG activity and cursor movement are marked with black arrows in *A*. *C*: EMG activity within the time window used for decoding movement direction (computed over all repetitions for each subject). Percentage changes from baseline were calculated for EMG activity and averaged for all trials. Error bars represent standard deviation. For overt movement, an increase in EMG activity will precede actual cursor movement, as governed by dynamics of limb movement. Temporally, EMG activity significantly overlaps with cortical activity representing intended movement as shown by previous animal studies (Moran and Schwartz 1999). Because we used time windows immediately before overt movement onset to decode intended movement direction, it is inevitable that there is an increase in EMG activity within those time windows as shown in *A* and *C*. Nevertheless, our data showed that there is no increase in EMG activity during imagined movement (*B* and *C*), which suggests that paralyzed patients could generate cortical activity encoding intended movement direction for brain-machine interface applications.



Supplemental Figure 2. Temporal dynamics of MEG sensor activity modulation by movement direction. The sensor-space MANOVA test was performed in 10-ms intervals to characterize movement modulation of magnetic activity, and the resulting F-statistics were plotted as a function of time for overt non-delayed (A), imagined non-delayed (B), overt delayed (C), and imagined delayed (D). For each plot, the thin gray lines represent the F-statistics for individual subjects, and the thick black lines represent F-statistics averaged over all subjects. Time 0 corresponds to movement onset for overt movement and target onset for imagined movement. Note that the MANOVA curves were used to find the optimal decoding window for the paper, so it is natural that the shape of the traces above are similar to the decoding accuracies shown in Figure 4.



Supplemental Figure 3. Comparison between cortical areas that are active and cortical areas that are tuned to intended movement direction. *A & B*: This is duplicated from Figure 5, which shows cortical areas that are active during overt and imagined movement, respectively. Color represents cortical source current amplitude, and the unit of the color bar is nAm. Hot color (yellow) indicates active areas. *C & D*: This is duplicated from Figure 6, which shows cortical areas that are tuned to movement directions during overt and imagined movement, respectively. Color represents F-statistic values. Hot color (yellow) indicates that activity of certain cortical area is significantly modulated by intended movement direction.



Supplemental Figure 4. Number of principal components that explain 99% of the variability in the sources. For each plot, the thin gray lines represent the number of components for individual subjects (non-delayed task), and the thick black lines represent the average over all subjects. Time 0 corresponds to movement onset for overt movement (A) and target onset for imagined movement (B). Note that the number of components needed to account for 99% of the variability in the sources reaches its minimum at the time of the movement for the overt task, and approximately at the time when the cursor starts moving in the imagined task.