De novo brain-computer interfacing deforms manifold of populational neural activity patterns in human cerebral cortex

Authors

Seitaro Iwama[1, 2], Zhang Yichi[1], Junichi Ushiba[3*]

Affiliations

[1]School of Fundamental Science and Technology, Graduate School of Keio University, Kanagawa, Japan; [2]Japan Society for the Promotion of Science, Tokyo, Japan; [3]Department of Biosciences and informatics, Faculty of Science and Technology, Keio University, Kanagawa, Japan

***Correspondence**

Junichi Ushiba, Ph.D.

Department of Biosciences and Informatics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kouhoku-ku, Yokohama, Kanagawa 223-8522, Japan.

Tel/Fax: +81-45-566-1678.

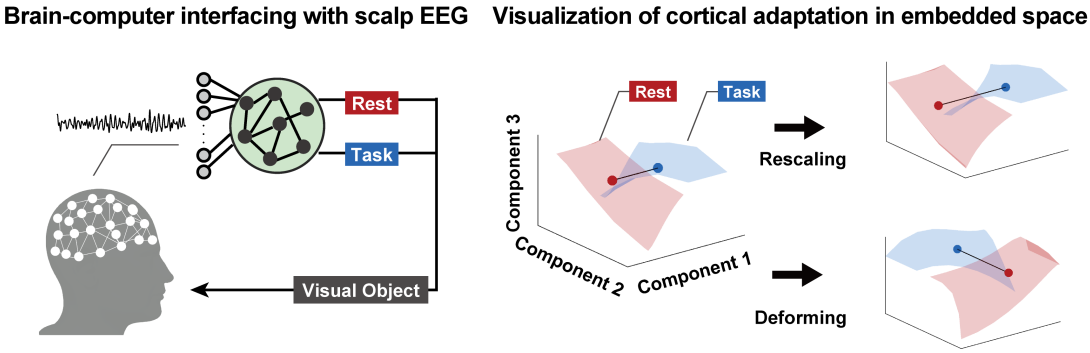
Email: ushiba@bio.keio.ac.jp

Abstract

Human brains are capable of modulating innate activities to adapt to novel environmental stimuli; for sensorimotor cortices (SM1) this means acquisition of a rich repertoire of motor behaviors. We investigated the adaptability of human SM1 motor control by analyzing net neural population activity during the learning of brain-computer interface (BCI) operations. We found systematic interactions between the neural manifold of cortical population activities and BCI classifiers; the neural manifold was stretched by rescaling motor-related features of electroencephalograms with model-based fixed classifiers, but not with adaptive classifiers that were constantly recalibrated to user activity. Moreover, operation of a BCI based on a *de novo* classifier with a fixed decision boundary based on biologically unnatural features, deformed the neural manifold to be orthogonal to the boundary. These principles of neural adaptation at a macroscopic level may underlie the ability of humans to learn wide-ranging behavioral repertoires and adapt to novel environments.

Keywords

brain-computer interface (BCI), nonlinear dimensionality reduction, sensorimotor activity, *de novo* learning, neural plasticity



1 Introduction

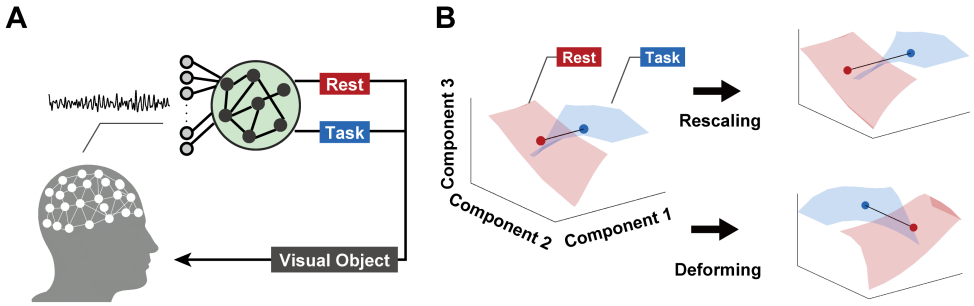
Neural plasticity underlies behavioral adaptation to the external environment by changing properties of neural circuitries involved in, for example, dexterous motor behaviors, such as sports, musical performance, tool-use, or brain-computer interface (BCI) operations (Imamizu et al., 2000; Nudo et al., 1996; Quallo et al., 2009). The adaptation processes to achieve purposeful physical movement have been examined by electrophysiology, neuroimaging, and behavioral approaches (Karni et al., 1995; Kleim et al., 2004; Kording et al., 2007; Nudo et al., 1996; Shadmehr & Mussa-Ivaldi, 1994).

In sensorimotor studies leveraging neural activity recordings, local neuronal circuitries display repertoires of firing patterns that reliably represent ongoing behavior (Gallego et al., 2018; Shenoy & Kao, 2021). This representation of covariance structure has been referred to as the neural manifold, and intriguing findings suggest that the brain is capable of rapidly learning patterns of spike activities inside the manifold but not those outside of it (Sadtler et al., 2014). These constraints to learning, which are putatively due to the microscopic configuration of neurons, illustrate realistic behavior as well as the BCI control on which neural activities and behavioral consequences are directly mapped.

While a neural manifold describes the constraints on the ensemble of local neural activities in which hundreds of neurons are implicated, what is less investigated are the constraints on the macroscopic neural system. Because the brain exerts information processing via not only local circuitry but also the inter-regional coupling by which macroscopic neural populations selectively communicate (Bassett et al., 2015), those implicated in coherent communication might also be constrained similarly to the local circuitries (Fries, 2005, 2015).

To characterize the constraints on cortical population activities during adaptation, we used BCIs based on scalp electroencephalograms (EEG) with a variety of incorporated classifiers (Figure 1).

Figure 1



Users attempted to move a virtual object using mental actions that modulated EEG signals. For each user, one of three classifiers determined the movement of objects based on a different set of rules. The model-based classifier required voluntary attenuation of sensorimotor rhythm (SMR) power derived from sensorimotor cortex (SM1). This fixed BCI operation rule is consistent with physiological findings, as the attenuation of SMR reflects SM1 excitability (Naros et al., 2019; Pfurtscheller & Lopes Da Silva, 1999; Takemi et al., 2013), as well as functional coupling among sensorimotor-related regions (Hayashi et al., 2020; Schulz et al., 2014; Tomassini et al., 2020; Wander et al., 2013). The adaptive classifier based on machine learning algorithms was configured based on recent whole-head EEG activity patterns to achieve maximum BCI controllability. This data-driven classifier configuration entails adaptive weighting on signaling features implicated in not only sensorimotor, but also attentional or cognitive functions in which the front-parietal network is implicated (Corsi et al., 2020). Lastly, the *de novo* classifier had a fixed configuration based on a biologically unnatural feature – desynchronized alpha oscillations derived from parietal regions. Due to the absence of prior knowledge to control this feature, users were encouraged to explore mental actions to control a visual object in the given BCI framework (Fujisawa et al., 2019).

As the decision boundaries between resting and motor attempts for each of the three classifiers (classifier plane) differed in their configurations, cortical adaptation processes were investigated by t-distributed stochastic neighbor embedding (t-SNE) algorithms, a nonlinear dimensionality reduction to visualize the distinct geometric changes of a whole-head EEG signal during operating/learning BCI tasks (Van Der Maaten & Hinton, 2008).

2 Results

2.1 Score acquisition during brain-computer interfacing

Twenty-one participants operated BCIs with one of three randomly allocated classifiers that provided scores contingent on BCI performance (Figure 2, Figure 2–supplement 1, 2). While BCI performance scores from the model-based and adaptive classifier generally increased over sessions, those for the *de novo* classifier did not. Statistical tests for coefficient of linear regression acquired from each participant revealed significant differences from zero for BCIs based on the model-based and adaptive classifiers (model-based: $p = 0.0078$, adaptive: $p =$

0.023, *de novo*: p = 0.055, Wilcoxon rank-sum test, FDR corrected). Note that direct comparison of the coefficients among classifiers is not possible because scores from each classifier were computed based on different classifiers.

Figure 2

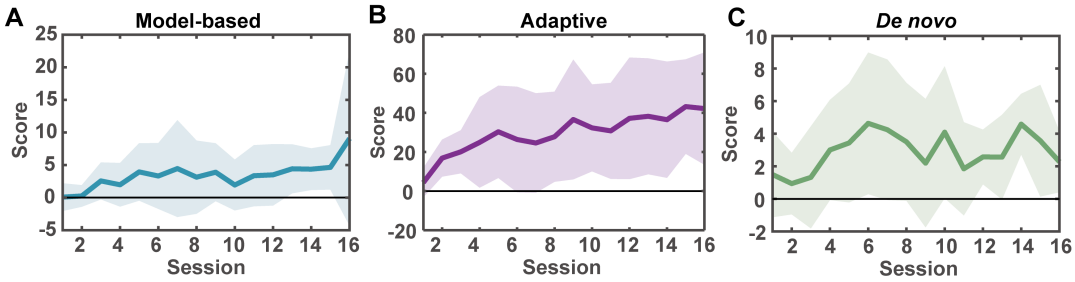


Figure 2 Supplement 1

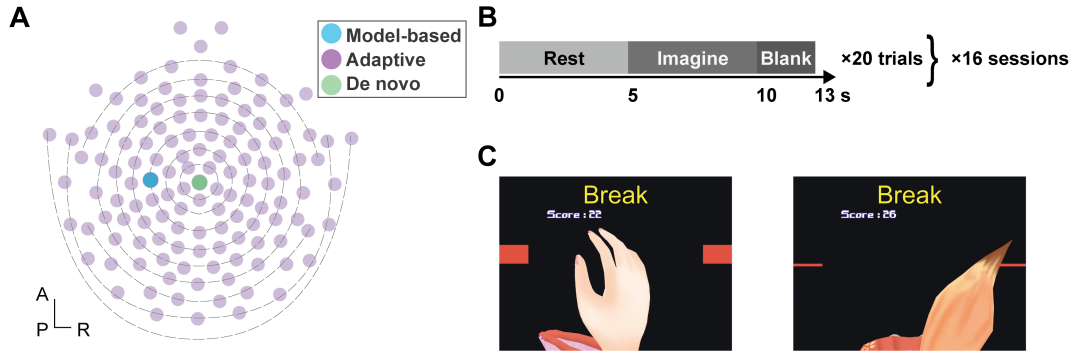
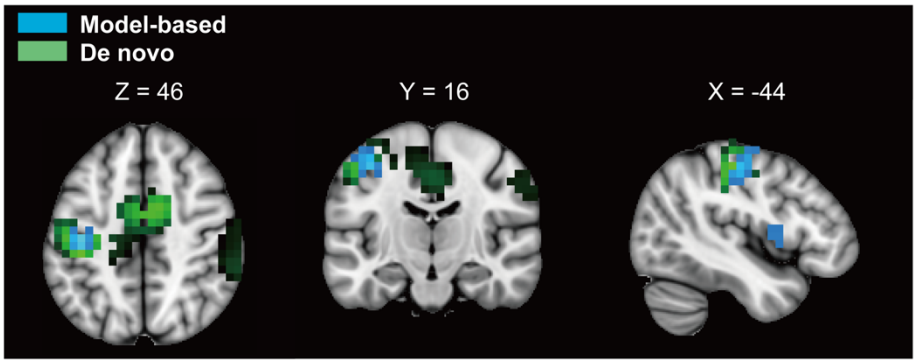


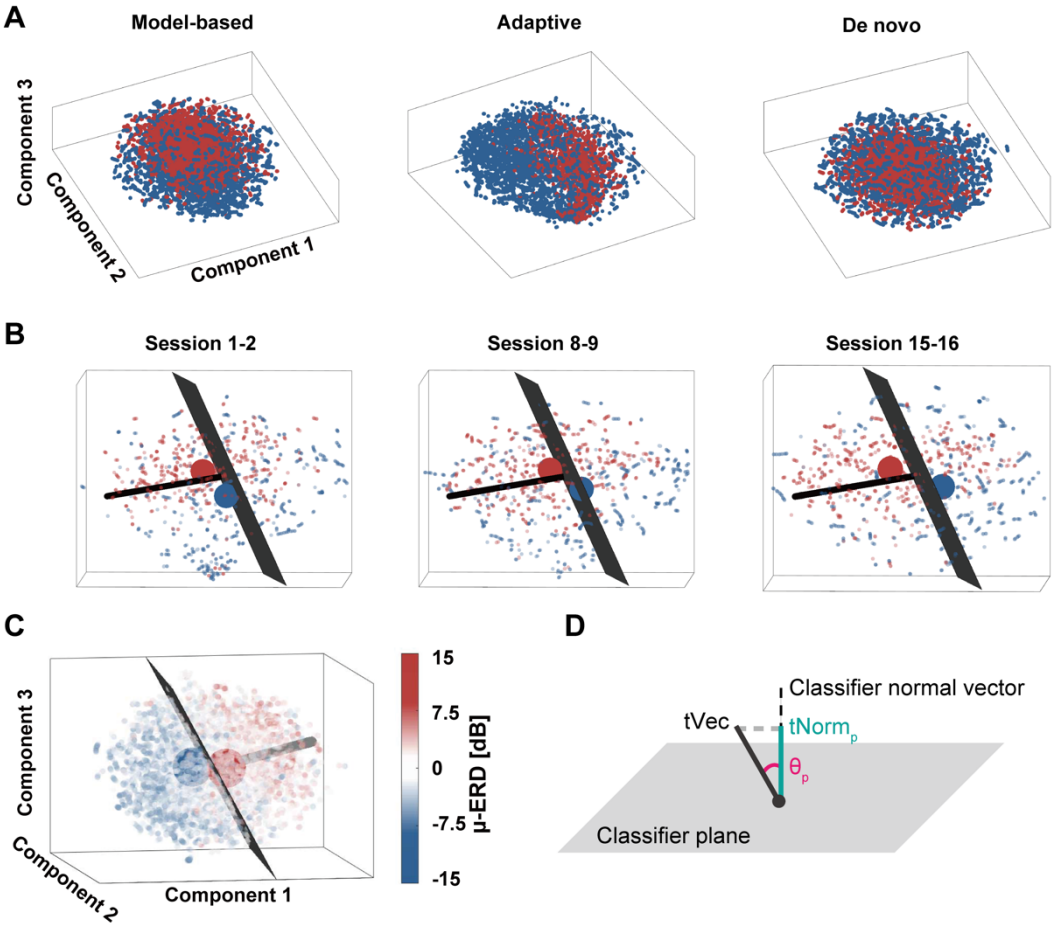
Figure 2 Supplement 2



2.2 Quantification of cortical adaptation process to classifier's separating plane

To examine differences in cortical adaptation processes, we next investigated changes in whole-head EEG signals (Figure 3A). Using band-power features as a representation of brain state, all data acquired from a single experiment were subjected to the t-SNE algorithm to evaluate geometric relationships among two brain states (i.e., resting and attempted movement) and the classifier plane in the embedded space.

Figure 3



An example of data from the model-based classifier BCI is shown in Figure 3B. As the participant performed the BCI operation, data during attempted movement (blue points) moved across the classifier plane, where the sign of relative SMR

power flips (Figure 3C). In this case, the defined metrics $tNorm_p$ and θ_p (Figure 3D) increased and decreased, respectively.

Figure 4A depicts changes in $tNorm_p$ and θ_p between the first and last four sessions in the experiment. For participants trained with the model-based classifier, $tNorm_p$ values significantly increased (p = 0.016, d = 0.71, two-tailed Wilcoxon signed-rank test) and the change was specific to participants who operated with model-based classifiers (Figure 4–supplement 1A, p = 0.81, 0.047). At the same time, θ_p values decreased significantly for participants trained with both the model-based and *de novo* classifiers (p = 0.016, d = 0.77, p = 0.016, d = 1.0, respectively, Figure 4–supplement 1B), but not with the adaptive classifiers (Figure. 4C; p = 0.58).

Figure 4

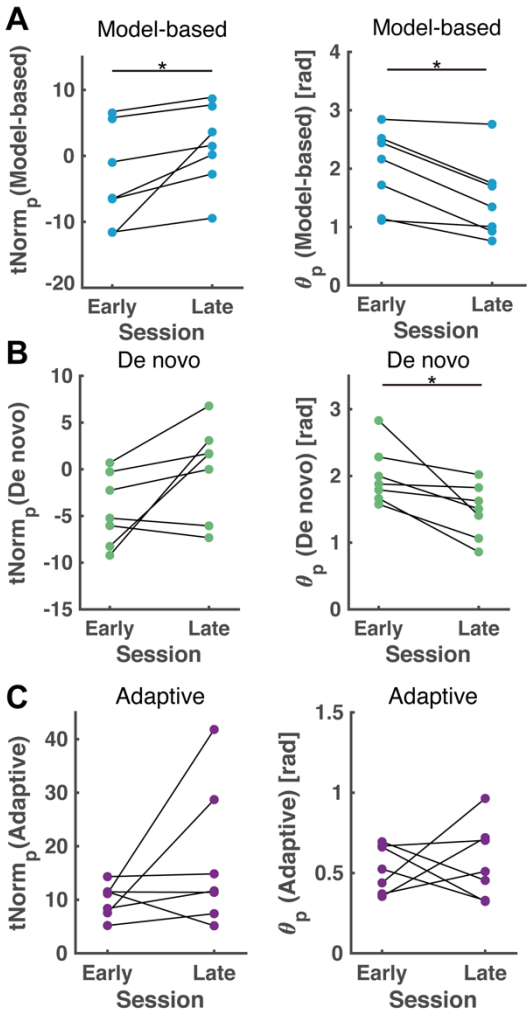


Figure 4 Supplement 1

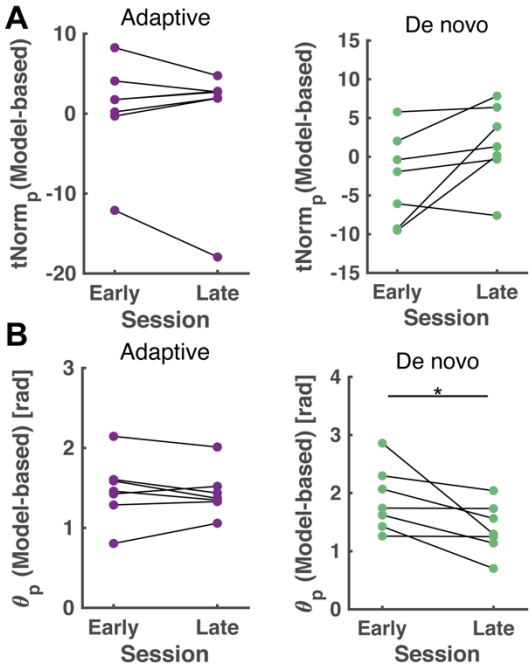
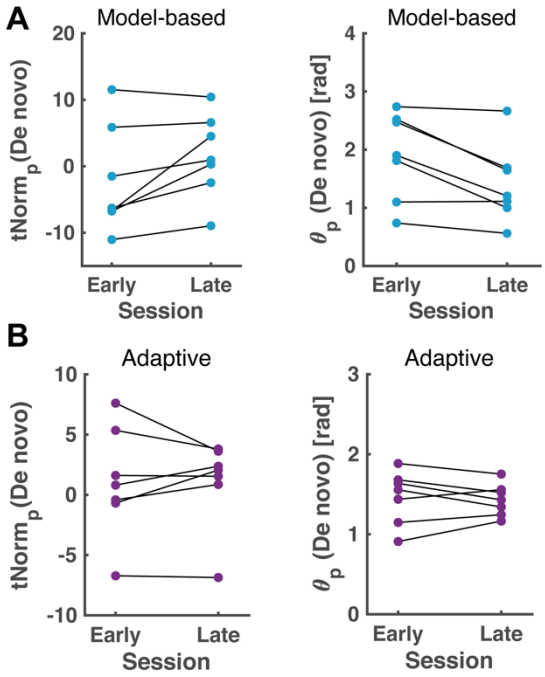
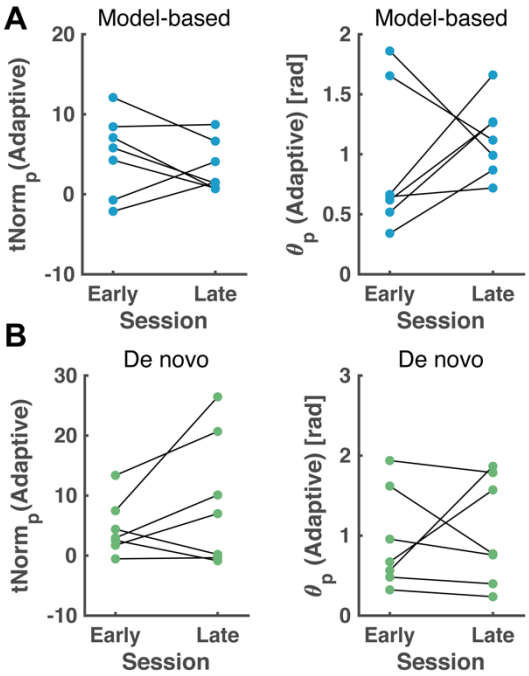


Figure 4 Supplement 2



156 **Figure 4 Supplement 3**



157

158 The identical evaluation was conducted for the *de novo* classifier plane.
159 Figure 4B depicts changes in $tNorm_p$ and θ_p against the *de novo* classifier.
160 While no significant differences were confirmed for $tNorm_p$ values over
161 sessions (p = 0.078), θ_p values decreased significantly (p = 0.016, d = 1.3).
162 Neither $tNorm_p$ nor θ_p changed with respect to the other two classifiers
163 (Figure 4–supplement 2, model-based: p = 0.047, 0.031, adaptive: p = 1, 0.58,
164 $tNorm_p$ and θ_p, respectively).

165 As the classifier planes changed from one session to the next for the adaptive
166 classifiers trained with the data from the previous sessions, each metric was
167 calculated against the classifier plane determined with the dataset from the
168 previous session. No significant differences were confirmed for comparison
169 between the early and late period for the adaptive classifier (Figure 4C,
170 supplement 3).

171

172 **3 Discussion**

173 In the present study, participants performed BCI operations with one of three
174 classifiers: model-based, adaptive, or *de novo*. Each classifier elicited a different
175 cortical adaptation process consistent with their characteristics. t-SNE analyses
176 in embedded space revealed increases in $tNorm_p$ for the model-based classifier,

177 indicating rescaling of the neural manifold with respect to the axes orthogonal to
178 the fixed decision boundary. Meanwhile, changes in population activities were not
179 induced by the adaptive classifiers; decreases in θ_p indicated that the manifold
180 was deformed, resulting in a reconfiguration orthogonal to its classifier plane by
181 the *de novo* classifier that was based on biologically unnatural features.
182

183 **3.1 Tuning classifiers to a brain induced by adaptive algorithm**
184 Both the model-based and adaptive classifiers elicited short-term learning of the
185 BCI operations as evidenced by the increases in performance scores; however,
186 these two processes were distinct from one another. While model-based
187 classifiers elicited changes in $tNorm_p$ and θ_p, the adaptive classifiers did not.
188 Such a difference might be attributed to the design of the classifiers, as the mental
189 actions that users were instructed to perform were identical. During BCI
190 operations with a constant classifier plane, participants honed their abstract
191 ability to control sensorimotor activity by minimizing error between the current
192 classified result and their intended mental action; however, in the case of the
193 adaptive classifiers, adaptation of users to the classifier was putatively limited
194 due to the session-by-session recalibration.
195 Despite the absence of cortical adaptation to the classifier plane for users of
196 the adaptive classifiers, performance scores did increase incrementally
197 throughout the experiment. Accordingly, we can only posit that the adaptation of
198 classifiers to users systematically progressed across sessions. It should be noted
199 that implementing the adaptive algorithm might induce suboptimal results when
200 the objective of the BCI operation is the induction of a specific neural activity,
201 such as changes in excitability, activity patterns, or connectivity of targeted
202 regions (Ramot et al., 2017; Ruddy et al., 2018; Shibata et al., 2011).
203

204 **3.2 Cortical adaptation process during *de novo* brain-computer interfacing**
205 Although significant increases in performance scores and $tNorm_p$ were not
206 confirmed for the *de novo* classifiers, cortical adaptations towards the classifier
207 plane were partly observed, as evidenced by the decreases in θ_p. The *de novo*
208 task was defined as one that participants work on to improve their performance
209 without any prior knowledge or strategy (Choi et al., 2020; Fujisawa et al., 2019;
210 Radhakrishnan et al., 2008; Telgen et al., 2014). To achieve this during brain-
211 computer interfacing, neurofeedback was provided via an illustrated tail. Because
212 movement of a tail is not inherent for humans, participants were instructed to

explore possible mental actions that might be suitable for operation. As such an exploratory strategy might require more extensive training than recalibrating the existing control configuration, performance scores did not tend to progress within a single-day experiment (Choi et al., 2020; Telgen et al., 2014).

The neural adaptation process was visualized via the t-SNE-based analysis. Deforming effects, that is rotational changes in the geometric relationship of two brain states towards the classifier plane, were confirmed in participants using the *de novo* classifier. However, the absence of a significant scaling effect suggested that the dissection of the two conditions (resting and motor imagery) did not systematically progress; this observation might reflect the reassociation of existing activity patterns to adapt to the BCI classifier by exploring a strategy to control the object. The result is consistent with the time course of the performance score and possible necessity of multi-day training to affect substantial behavioral improvement in *de novo* learning (Choi et al., 2020; Fujisawa et al., 2019). Although the flexibility of the human brain enabled partial adaptation to the *de novo* classifier planes, the adaptive classifier did not elicit brain-side adaptation. These findings collectively suggest that fixation of the classifier plane is an essential element for inducing neural plasticity via a brain-computer interaction based on macroscopic neural population activity.

4 Material and Methods

4.1 Participants

Twenty-one neurologically healthy adults (9 females, 12 males, mean age: 22.6 ± 3.23) who had never operated a BCI participated in this experiment. The appropriate sample size for this study was determined by an a-priori power analysis ($\alpha = 0.05$, $1-\beta = 0.8$, two-sided Wilcoxon signed-rank tests) focusing on the deforming effect induced by *de novo* BCI. The statistical package G*Power 3 (Faul et al, 2007) was used to estimate the sample size that shows large Cohen's $d = 0.90$ reported in the previous EEG-based neurofeedback literatures (Hayashi et al., 2020; Soekadar et al., 2015). We calculated that 7 participants were needed.

All participants had normal or corrected-to-normal vision and were asked to provide written informed consent before participating in the experiment. This study was conducted according to the ethics of the Declaration of Helsinki. The

248 experimental protocol was approved by the ethical committee of the Faculty of
249 Science and Technology, Keio University (Approval Number: 2020-36, 31-23).
250
251 **4.2 Experimental setup**
252 Participants were seated on a comfortable chair in a quiet room. A display was
253 placed about one meter in front of the chair to provide task instructions and visual
254 feedback from BCIs.
255 EEG signals during the experiment were acquired with a 128-channel
256 HydroCel Geodesic Sensor Net (HCGSN, EGI, Eugene, OR, USA.). The layout
257 of channels followed the international 10-10 electrode positions shown in
258 Figure 2−supplement 1A (Luu & Ferree, 2005). The reference channel was set to
259 Cz. The impedance of all channels was maintained below 50 kΩ throughout the
260 experiment. The EEG data were collected with a sampling rate of 1 kHz and
261 transmitted via the Ethernet switch Gigabit Web Smart Switch (Black Box,
262 Pennsylvania, USA) to EEG recording software Net Station 5.2 manufactured by
263 EGI and MATLAB R2019a (The Mathworks, Inc, Massachusetts, USA).

264 **4.3 Online processing of EEG signals**
265 Analytical pipelines for online signal processing were implemented with a
266 combination of MATLAB and Unity (Ver. 2019.2.4f1, Unity Technologies, USA).
267 While the real-time status of brain states was determined from EEG signals
268 processed by a custom MATLAB script, Unity presented the visual neurofeedback
269 objects. EEG signals were processed with a 1651-point, minimum-phase, FIR 8-
270 30 Hz bandpass temporal filter and then processed with one of the three types of
271 BCI classifiers. Online processed EEG signals were used to classify brain-states
272 as either presence or absence of attempted movement with one of the three types
273 of classifiers: model-based, adaptive, or *de novo*. Each classifier was designed
274 with different rules, and electrodes of interest were defined as shown in Figure 2−
275 supplement 1A.
276 The model-based classifier was constructed based on those used in common
277 SMR-based BCIs (Buch et al., 2008; Kraus et al., 2016). EEG signals around the
278 left SM1 (i.e., channel C3) were only used to detect the attempted movement,
279 because accumulated evidence suggests that event-related desynchronization of
280 SMR (SMR-ERD) contralateral to the hand that attempted to move reflects the
281 excitability of SM1 (Hummel et al., 2002; Naros et al., 2019; Takemi et al., 2013).
282 In online processing, a large Laplacian filter was applied to EEG signals from
283 channel C3 to extract sensorimotor activity (McFarland et al., 1997; Tsuchimoto

et al., 2021). Subsequently, the band power of SMR (SMR-power; 8-13 Hz) was extracted by Fourier transform with a 1-s window and Hamming window function. The magnitude of SMR-ERD [dB] was computed from the obtained SMR-power with the following formula:

$$\mathrm{ERD}(t) = -10\log_{10}(\mathrm{P}(t)\,/\mathrm{P}_{Ref})$$

where $\mathrm{P}(t)$ denotes the power of interest, here the SMR-power, at time point t, and P_{Ref} denotes the reference power (Pfurtscheller & Lopes Da Silva, 1999). The reference power was calculated from the middle 3-s period of "Rest" time from the previous trial. The online calculated magnitude of SMR-ERD was then used as the index of neurofeedback for the model-based classifier. Movements of the illustrated hand in the display and performance scores were defined to be linearly related to the SMR-ERD value in the range of 0 to 10 dB.

The adaptive classifier was constructed using whole-head scalp EEG signals based on a common spatial pattern (CSP) algorithm and a support vector machine (SVM) (Blankertz et al., 2007). CSP components were extracted to maximize the separability of the two conditions Rest and Imagine, and were quickly trained at the end of each session to adapt to the current activity patterns of users. Specifically, the CSP was generated from the spatial covariance matrices of all EEG electrodes to find linear combinations of electrodes to form spatial filters that maximized the variance difference between the two conditions. The corresponding variances of spatially filtered EEG data were then divided into time windows and log-transformed to transform their distribution into a normal distribution. The SVM classifier was constructed to perform a binary classification of the two conditions. The posterior probability for a data point classified as presence of motor attempt was used as an index for neurofeedback; the index for the adaptive classifier was defined to be linearly related to the posterior possibility in the range of 50% to 100%. Note that the rules for object movement were identical to those of the model-based classifier, only the feedback was different.

Lastly, the *de novo* classifier had a fixed classifier plane as did the model-based classifier; however, its characteristics were biologically unnatural; the *de novo* classifier was based on EEG signals around the parietal region (i.e., channel Cz) that are associated with attentional features but not with sensorimotor activity (Benedek et al., 2014; Misselhorn et al., 2019). During the BCI task, users

attempted to move their body or a visual object on the display; however, spectral power in the alpha-band (8-13 Hz) was increased by the motor attempt of moving the feet or by internal attention at the targeted channel (Benedek et al., 2014; Pfurtscheller et al., 2006). Such intrinsic responses did not contribute to the BCI operation, as the *de novo* classifier discriminated motor attempts with ERD values (i.e., power attenuation) in the alpha-band from the Cz channel, calculated with the procedure identical to that from channel C3 in the model-based classifier. Online computed ERD magnitude was exploited to decode the absence/presence of attempted movement and index for neurofeedback. Note that the rules for object movement and for obtaining scores were identical to those in the other two types of classifiers.

4.4 Experimental procedure

Participants underwent 16 BCI operation sessions, each consisting of 20 trials. All experimental procedures were conducted within 2 hours to guarantee the reversibility of any potentially unnaturally induced neural plasticity and to investigate the initial phase of learning to operate the BCIs. After every two sessions, participants were given a break of up to 5 min. Participants were randomly allocated to one of the three classifiers without knowledge of their existence or configuration and used the allocated type of classifier throughout the entire experiment.

A trial began with a 5-s “Rest” period. This was followed by a 5-s “Imagine” and a 3-s “Break” period (Figure 2 – supplement 1B). During the “Rest” period, participants were instructed to relax without having any specific thoughts and with opened eyes. In the “Imagine” period, participants were instructed by the experimenter to perform motor imagery tasks based on the allocated classifier. Participants with the model-based and adaptive classifiers were instructed to imagine extending the right-hand throughout the experiment, matching the imagined movement with the object on display. Participants with the *de novo* classifier were instructed to imagine moving a tail, also matching the movement of the object on display. As tail moving is not intuitive for humans, at the beginning of the session, participants were encouraged to exploratorily find a strategy that achieved the best control of the BCI. The strategy adopted in each session was freely determined by each participant, but they were instructed to try to use the same strategy throughout one session to acquire sufficient data during a specific strategy.

The performance of each trial was quantified by a score, which participants were encouraged to make as high as possible. Scores were determined by the predicted presence/absence of the attempted movement. Both the absence of attempted movements during “Rest” periods and the presence of attempted movements during “Imagine” periods resulted in higher scores, while scores were reduced if movements contrary these predictions were detected. The changing rates of these scores were pertinent to the metrics used for feedback by each classifier and were regulated linearly to fit the score range from minus one hundred to plus one hundred. After each session, participants were asked to verbally describe the strategies they had adopted.

For the adaptive classifier, the CSP-SVM model was re-trained with “Rest” and “Imagine” period data from the previous session to use in the next session. Note that the first session of the adaptive classifier task was identical to that of the model-based one, so as to collect a dataset for constructing the adaptive classifier. Detailed procedures for classifier training are described in the supplementary materials.

4.5 Evaluation of BCI performance

Online-calculated scores were subjected to linear regression analysis (Gruzelier, 2014; Kober et al., 2018; Witte et al., 2018). The score obtained during a given session was used as a dependent variable and session number was used as a predictor valuable. If scores increased during the experiment, the regression coefficient for the predictor valuable was positive. To test whether the obtained regression coefficients were significantly different from zero, they were subjected to a group-by-group Wilcoxon rank-sum test with a false discovery rate correction (Benjamini-Hochberg method; Benjamini & Hochberg, 1995).

4.6 Offline EEG preprocession

The recorded EEG signals were first preprocessed with EEGLAB (Delorme & Makeig, 2004) to reject artifacts and enhance the computational efficiency via downsampling (Bigdely-Shamlo et al., 2015). The raw EEG data were filtered with a zero-phase 1-45 Hz FIR bandpass filter and down sampled to 100 Hz. Channels classified as “Bad” by the EEGLAB plugin: Christian's clean_rawdata (Bigdely-Shamlo et al., 2015) were removed from further analysis. The removed channels were interpolated spherically to minimize a potential bias when re-referencing the electrodes to a common average reference. Subsequently, large-amplitude

artifacts caused by blinking or head displacement were removed via Artifact Subspace Reconstruction (Kothe & Makeig, 2013). The electrodes were then re-referenced to the common average reference to extract activity specific to the electrodes (McFarland et al., 1997).

The continuous EEG data were then segmented into trials to evaluate the middle 8-s periods of the online BCI training trials (i.e., the last 4 s of the "Rest" period and the first 4 s of the "Imagine" period). To obtain the independent EEG components of the segmented dataset, we used adaptive mixture independent component analysis (AMICA; Palmer et al., 2011). Finally, an automatic artifact rejection was applied using ICLabel that distinguished brain–originated EEG components from artifacts induced by eye, muscle, heart, line noise, and channel noises (Pion-Tonachini et al., 2019).

To investigate cortical adaptation processes during brain-computer interfacing, the band-power features were used as a raw-vector that represents instantaneous overall brain state. Computed band-power from each EEG channel was subdivided into five functionally distinct frequency bands (Delta: 1-4 Hz, Theta: 4-8 Hz, Alpha: 8-13 Hz, Beta: 13-31 Hz, Gamma: 31-45 Hz; Hayashi et al., 2019). The averaged band-power was log-transformed and normalized to the z-score in a trial-by-trial manner to cancel base-line drifting. Thereby, the original number of dimensions of the feature vector D was $D = 129 \times 5 = 645$. The calculated band-power signals in the alpha-band also were subjected to cortical source estimation (See also supplementary materials).

4.7 Feature extraction of EEG-dataset using t-SNE algorithm

The preprocessed EEG dataset (645×11520 matrix) was subjected to a subject-by-subject t-SNE analysis, which converted the pairwise distances between data points in the original feature space to conditional probabilities (Van Der Maaten & Hinton, 2008). Mathematical details of t-SNE are further described in the Supplementary Materials and briefly here. First, the conditional probability that the data points x_i and x_j are neighbors was calculated from the pairwise distances of input data. Then, to maintain the probabilities in the original feature space in the embedded space, the Kullback-Leibler divergence representing the distance between the conditional probability in the original and embedded space was minimized via optimization. In this study, the number of dimensions of EEG features was reduced to three via a Barnes-Hut variation of t-SNE (Van Der Maaten et al., 2014) to speed up the computation. Perplexity, a hyperparameter

of the t-SNE algorithm, was set to 20, which was determined empirically with a parameter search of past EEG data for best separation between the “Rest” and “Imagine” periods. The hyperparameter was fixed across participants throughout the study after the determination. After applying t-SNE, the dimensionality-reduced datasets were subjected to visualization and a similarity analysis.

4.8 t-SNE-based dimensionality reduction and quantitative analysis in embedded space

Feature extraction using dimensionality reduction is popularly conducted for high-dimensional neural data across modalities (Cunningham & Yu, 2014; Lord et al., 2019). The t-SNE algorithm we adopted has advantages for geometric evaluation, as it preserves original distances in the embedded space. Here, to conduct quantitative analysis beyond its general purpose for data visualization, we employed metrics that do not violate the assumption of the t-SNE algorithm. Feature extraction techniques such as ICA, principal component analysis, or factor analysis display weighted maps of extracted components so that they can be applied to newly acquired data, whereas the t-SNE algorithm does not. To enable interpretation of the dataset with reduced components, in the present study, entire datasets from individual participants were subjected to t-SNE analysis at once.

Because t-SNE unfolds the nonlinear structure of a given dataset, the linear distance in the embedded space can be interpreted as an approximation of geometric distance in the original space. It illustrates how different one brain activity pattern is from another. It should be noted however that to properly interpret the results (1) distance scales in the embedded space were rearranged and were variable across iterations of t-SNE, (2) distance scales in different clusters might have differed, (3) direct comparisons of distances between clusters were not acceptable because distances within two clusters were arbitrary. To deal with the above concerns, two approaches were adopted: (1) data points were bridged to prevent the formation of multiple clusters, and (2) statistical distances, namely Hotelling's t-squared statistical values, were used instead of Euclidean metrics.

Because distances between nearby points are well preserved in embedded space, the distance scale of distant points were kept similar for enough data points, which acts as a bridge and prevents the formation of sparse multiple clusters. We also adopted the concept of “short-circuiting” (Lee & Verleysen,

2005) by constructing the feature vectors with overlapped time-windows so that points were smoothly connected. Thus, distances from point to point shared the same scale across all points (i.e., only one cluster was generated in embedded space).

Hotelling's t-squared statistic was adopted as the distance metrics between two group of points (Hotelling, 1992). Assume x and y are two groups of points lying in a *p*-dimensional space, n_x and n_y are the numbers of points, $\bar{x}$ and $\bar{y}$ are the sample means, and $\hat{\Sigma}_x$ and $\hat{\Sigma}_y$ are the respective sample covariance matrices. The Hotelling's t-squared statistic was calculated as:

$$\boldsymbol{t^2 = \frac{n_x n_y}{n_x + n_y}(\bar{x} - \bar{y})' \hat{\Sigma}^{-1}(\bar{x} - \bar{y})}$$

$$\hat{\Sigma} = \frac{(n_x - 1)\hat{\Sigma}_x + (n_y - 1)\hat{\Sigma}_y}{n_x + n_y - 2}$$

Hotelling's t-squared statistic is suitable for measurements of statistical distance in the t-SNE-embedded space, as they were invariant to the distance scale. The distribution of t^2 follows an *F*-distribution:

$$t^2 \sim \frac{p(n_x + n_y - 2)}{n_x + n_y - p - 1} F_{p, n_x + n_y - 1 - p}$$

To normalize the distribution, the square root of t^2 was defined as $tNorm$ and was used as the distance measurement in subsequent analyses:

$$t\mathrm{Norm} = \sqrt{t^2}$$

The vector representing the directional relationship between two classes was defined as $tVec$:

$$t\mathrm{Vec} = t\mathrm{Norm} \cdot \frac{\bar{x} - \bar{y}}{\|\bar{x} - \bar{y}\|}$$

Data points were divided into two classes: “Rest” and “Imagine” according to their relative times in the trials. $tNorm$ and $tVec$ were calculated for these two conditions.

4.9 Classifier plane

To investigate the influence of BCI classifiers on the cortical adaptation in the t-SNE-embedded space, the classifier plane and classifier normal vector were linearly projected into the embedded space (See Figure 3C). The classifier vector $V = [v_1, v_2, v_3]^T$ was calculated as follows, where T denotes a matrix transpose.

$$X = \begin{pmatrix} 1 & \\ \vdots & Y \\ 1 & \end{pmatrix}, \qquad \begin{pmatrix} d \\ \vec{v} \end{pmatrix} = X^{-1}\, X^T P \quad (2.9)$$

$$\boldsymbol{V} = \vec{\boldsymbol{v}} / \|\vec{\boldsymbol{v}}\|$$

Then, the equation of the classifier plane is given as follows.

$$\boldsymbol{v_1 x + v_2 y + v_3 z + b = 0}$$

assuming $Y \in \mathbb{R}^{N\times 3}$ are the points in the 3D embedded space, $P \in \mathbb{R}^{N\times n}$ are the original features referred to by the classifier, where *N* is the number of points, and n is the number of features. $\boldsymbol{b}$ is the intercept corresponding to the decision boundary of the classifiers.

As is shown in Figure 3D, $tVec$ could be projected to the classifier vector to evaluate its positional relationship against the classifier. The lengths of projection on the classifier vector ($t\mathrm{Norm_p}$) and the angles between $tVec$ and the classifier vector and (θ_p) were calculated against that of the model-based classifier to directly compare the adaptation processes across classifiers as follows:

$$t\mathrm{Norm_p} = t\mathrm{Vec} \cdot V$$

$$\theta_p = \arccos \frac{t\mathrm{Vec} \cdot V}{\|t\mathrm{Vec}\|}$$

which represent the strength of scaling and deforming against the classifier plane, respectively.

4.10 EEG-similarity analysis

Geometry-based analysis was conducted in the embedded space, as positional relationships of the points reflected the similarities in the original space. The

transition process from one brain condition to another (i.e., absence to presence of attempted movement) was assessed by the spatial arrangement and separability of points from the “Rest” and “Imagine” periods in the t-SNE dimension. Emergence of the two temporal phenomena were defined as follows:

- Scaling: The separability of the two conditions (Rest and Imagine) increases with respect to a fixed axis. Scaling is interpreted as the enhancement of specific cortical activity patterns.
- Deforming: The relationship of positions in the two conditions changes direction. Deforming is interpreted as an alteration of a cortical activity pattern that is adopted.

To quantify the two distinct adaptation process, the following metrics were defined. Scaling and deforming between the i^{th} and j^{th} sessions were quantified by $tNorm_p$ and θ_p.

- Scaling: $\Delta t\mathrm{Norm}_p = t\mathrm{Norm}_p(i) - t\mathrm{Norm}_p(j)$
- Deforming: $\Delta\theta_p = \theta_p(i) - \theta_p(j)$

If adaptation progresses toward the targeted neural activity patterns required to control BCIs, the $t\mathrm{Norm}_{\mathrm{p}}$ values should be larger while those of θ_p should be smaller. Thus, the calculated values were subjected to the Wilcoxon sign-rank test to compare the differences between the first and last four sessions (early and late period, respectively). For adaptive classifiers, as the classifier plane was obtained from the 2nd session, the comparison was conducted from the 2nd session across all statistical tests. We then corrected the alpha-level with a Bonferroni correction.

Acknowledgements

This study was supported by the Keio Institute of Pure and Applied Sciences (KiPAS) research program, JSPS KAKENHI Grant Number 20H05923 (to J.U.) and JST, CREST Grant Number JPMJCR17A3 (to J.U.) including the AIP challenge program, Japan. We thank Yumiko Kakubari, Shoko Tonomoto and Aya Kamiya for their technical supports.

Competing interests

J.U. is a founder and Representative Director of the university startup company, Connect Inc. involved in the research, development, and sales of rehabilitation devices including brain-computer interfaces. He receives a salary from Connect Inc., and holds shares in Connect Inc. This company does not have any relationships with the device or setup used in the current study.

References

[1] Bassett, D. S., Yang, M., Wymbs, N. F., & Grafton, S. T. (2015). Learning-induced autonomy of sensorimotor systems. *Nature Neuroscience*, *18*(5), 744–751. https://doi.org/10.1038/nn.3993

[2] Benedek, M., Schickel, R. J., Jauk, E., Fink, A., & Neubauer, A. C. (2014). Alpha power increases in right parietal cortex reflects focused internal attention. *Neuropsychologia*, 56(1), 393–400. https://doi.org/10.1016/j.neuropsychologia.2014.02.010

[3] Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

[4] Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*, 1–19. https://doi.org/10.3389/fninf.2015.00016

[5] Blankertz, B., Dornhege, G., Krauledat, M., Müller, K. R., & Curio, G. (2007). The non-invasive Berlin Brain-Computer Interface: Fast acquisition of effective performance in untrained subjects. *NeuroImage*, *37*(2), 539–550. https://doi.org/10.1016/j.neuroimage.2007.01.051

[6] Buch, E., Weber, C., Cohen, L. G., Braun, C., Dimyan, M. A., Ard, T., Mellinger, J., Caria, A., Soekadar, S., Fourkas, A., & Birbaumer, N. (2008). Think to move: A neuromagnetic brain-computer interface (BCI) system for chronic stroke. *Stroke*, *39*(3), 910–917. https://doi.org/10.1161/STROKEAHA.107.505313

[7] Choi, Y., Shin, E. Y., & Kim, S. (2020). Spatiotemporal dissociation of fMRI activity in the caudate nucleus underlies human de novo motor skill learning. *Proceedings of the National Academy of Sciences of the United*

States of America, *117*(38), 23886–23897. https://doi.org/10.1073/pnas.2003963117

[8] Corsi, M. C., Chavez, M., Schwartz, D., George, N., Hugueville, L., Kahn, A. E., Dupont, S., Bassett, D. S., & De Vico Fallani, F. (2020). Functional disconnection of associative cortical areas predicts performance during BCI training. *NeuroImage*, *209*, 116500. https://doi.org/10.1016/j.neuroimage.2019.116500

[9] Cunningham, J. P., & Yu, B. M. (2014). Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*, *17*(11), 1500–1509. https://doi.org/10.1038/nn.3776

[10] Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *13*, 9–21.

[11] Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. https://doi.org/10.3758/BF03193146

[12] Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011

[13] Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, *88*(1), 220–235. https://doi.org/10.1016/j.neuron.2015.09.034

[14] Fujisawa, A., Kasuga, S., Suzuki, T., & Ushiba, J. (2019). Acquisition of a mental strategy to control a virtual tail via brain–computer interface. *Cognitive Neuroscience*, *10*(1), 30–43. https://doi.org/10.1080/17588928.2018.1426564

[15] Gallego, J. A., Perich, M. G., Naufel, S. N., Ethier, C., Solla, S. A., & Miller, L. E. (2018). Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nature Communications*, *9*(1), 1–13. https://doi.org/10.1038/s41467-018-06560-z

[16] Gruzelier, J. H. (2014). EEG-neurofeedback for optimising performance. III: A review of methodological and theoretical considerations. *Neuroscience and Biobehavioral Reviews*, *44*, 159–182. https://doi.org/10.1016/j.neubiorev.2014.03.015

[17] Hayashi, M., Mizuguchi, N., Tsuchimoto, S., & Ushiba, J. (2020). Neurofeedback of scalp bi-hemispheric EEG sensorimotor rhythm guides hemispheric activation of sensorimotor cortex in the targeted hemisphere. *NeuroImage*, *223*, 117298. https://doi.org/10.1016/j.neuroimage.2020.117298

[18] Hayashi, M., Tsuchimoto, S., Mizuguchi, N., Miyatake, M., Kasuga, S., & Ushiba, J. (2019). Two-stage regression of high-density scalp electroencephalograms visualizes force regulation signaling during muscle contraction. *Journal of Neural Engineering*, *16*(5), 056020. https://doi.org/10.1088/1741-2552/ab221a

[19] Hotelling, H. (1992). *The Generalization of Student's Ratio* (pp. 54–65). Springer, New York, NY. https://doi.org/10.1007/978-1-4612-0919-5_4

[20] Hummel, F., Andres, F., Altenmüller, E., Dichgans, J., & Gerloff, C. (2002). Inhibitory control of acquired motor programmes in the human brain. *Brain*, *125*(2), 404–420. https://doi.org/10.1093/brain/awf030

[21] Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Pütz, B., Yoshioka, T., & Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, *403*(6766), 192–195. https://doi.org/10.1038/35003194

[22] Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, *377*(6545), 155–158. https://doi.org/10.1038/377155a0

[23] Kleim, J. A., Hogg, T. M., VandenBerg, P. M., Cooper, N. R., Bruneau, R., & Remple, M. (2004). Cortical Synaptogenesis and Motor Map Reorganization Occur during Late, but Not Early, Phase of Motor Skill Learning. *Journal of Neuroscience*, *24*(3), 628–633. https://doi.org/10.1523/JNEUROSCI.3440-03.2004

[24] Kober, S. E., Witte, M., Grinschgl, S., Neuper, C., & Wood, G. (2018). Placebo hampers ability to self-regulate brain activity: A double-blind sham-controlled neurofeedback study. *NeuroImage*, *181*, 797–806. https://doi.org/10.1016/j.neuroimage.2018.07.025

[25] Kording, K. P., Tenenbaum, J. B., & Shadmehr, R. (2007). The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nature Neuroscience*, *10*(6), 779–786. https://doi.org/10.1038/nn1901

669 [26] Kothe, C. A., & Makeig, S. (2013). BCILAB: A platform for brain-computer
670 interface development. *Journal of Neural Engineering*, *10*(5), 56014–56031.
671 https://doi.org/10.1088/1741-2560/10/5/056014
672 [27] Kraus, D., Naros, G., Bauer, R., Leão, M. T., Ziemann, U., & Gharabaghi, A.
673 (2016). Brain-robot interface driven plasticity: Distributed modulation of
674 corticospinal excitability. *NeuroImage*, *125*, 522–532.
675 https://doi.org/10.1016/j.neuroimage.2015.09.074
676 [28] Lee, J. A., & Verleysen, M. (2005). Nonlinear dimensionality reduction of
677 data manifolds with essential loops. *Neurocomputing*, *67*, 29–53.
678 https://doi.org/10.1016/j.neucom.2004.11.042
679 [29] Lord, L. D., Expert, P., Atasoy, S., Roseman, L., Rapuano, K., Lambiotte,
680 R., Nutt, D. J., Deco, G., Carhart-Harris, R. L., Kringelbach, M. L., & Cabral,
681 J. (2019). Dynamical exploration of the repertoire of brain networks at rest is
682 modulated by psilocybin. *NeuroImage*, *199*, 127–142.
683 https://doi.org/10.1016/j.neuroimage.2019.05.060
684 [30] Luu, P., & Ferree, T. (2005). Determination of the HydroCel Geodesic
685 Sensor Nets' Average Electrode Positions and Their 10 – 10 International
686 Equivalents. *EGI Technical Note*, *October*, 1–11.
687 [31] McFarland, D. J., McCane, L. M., David, S. V., & Wolpaw, J. R. (1997).
688 Spatial filter selection for EEG-based communication.
689 *Electroencephalography and Clinical Neurophysiology*, *103*(3), 386–394.
690 https://doi.org/10.1016/S0013-4694(97)00022-2
691 [32] Misselhorn, J., Friese, U., & Engel, A. K. (2019). Frontal and parietal alpha
692 oscillations reflect attentional modulation of cross-modal matching. *Scientific*
693 *Reports*, 9(1), 1–11. https://doi.org/10.1038/s41598-019-41636-w
694 [33] Naros, G., Lehnertz, T., Leão, M. T., Ziemann, U., & Gharabaghi, A. (2019).
695 Brain State-dependent Gain Modulation of Corticospinal Output in the
696 Active Motor System. *Cerebral Cortex*, *00*, 1–11.
697 https://doi.org/10.1093/cercor/bhz093
698 [34] Nudo, R. J., Milliken, G. W., Jenkins, W. M., & Merzenich, M. M. (1996).
699 Use-dependent alterations of movement representations in primary motor
700 cortex of adult squirrel monkeys. *Journal of Neuroscience*, *16*(2), 785–807.
701 https://doi.org/10.1523/jneurosci.16-02-00785.1996
702 [35] Palmer, J., Kreutz-Delgado, K., & Makeig, S. (2011). AMICA: An Adaptive
703 Mixture of Independent Component Analyzers with Shared Components.
704 *San Diego, CA: Technical Report, Swartz Center for Computational*

Neuroscience, *January*, 1–15. http://sccn.ucsd.edu/~jason/amica_a.pdf%5Cnpapers2://publication/uuid/E6296FC1-7F6B-400C-85D0-3A292A27F710

[36] Pfurtscheller, G., Brunner, C., Schlögl, A., & Lopes da Silva, F. H. (2006). Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. NeuroImage, 31(1), 153–159. https://doi.org/10.1016/j.neuroimage.2005.12.003

[37] Pfurtscheller, G., & Lopes Da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857. https://doi.org/10.1016/S1388-2457(99)00141-8

[38] Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

[39] Quallo, M. M., Price, C. J., Ueno, K., Asamizuya, T., Cheng, K., Lemon, R. N., & Iriki, A. (2009). Gray and white matter changes associated with tool-use learning in macaque monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(43), 18379–18384. https://doi.org/10.1073/pnas.0909751106

[40] Radhakrishnan, S. M., Baker, S. N., & Jackson, A. (2008). Learning a novel myoelectric-controlled interface task. *Journal of Neurophysiology*, *100*(4), 2397–2408. https://doi.org/10.1152/jn.90614.2008

[41] Ramot, M., Kimmich, S., Gonzalez-Castillo, J., Roopchansingh, V., Popal, H., White, E., Gotts, S. J., & Martin, A. (2017). Direct modulation of aberrant brain network connectivity through real-time NeuroFeedback. *ELife*, *6*. https://doi.org/10.7554/eLife.28974

[42] Ruddy, K., Balsters, J., Mantini, D., Liu, Q., Kassraian-Fard, P., Enz, N., Mihelj, E., Subhash Chander, B., Soekadar, S. R., & Wenderoth, N. (2018). Neural activity related to volitional regulation of cortical excitability. *ELife*, *7*, 21. https://doi.org/10.7554/elife.40843

[43] Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., Yu, B. M., & Batista, A. P. (2014). Neural constraints on learning. *Nature*, *512*(7515), 423–426. https://doi.org/10.1038/nature13665

[44] Schulz, H., Übelacker, T., Keil, J., Müller, N., & Weisz, N. (2014). Now i am ready - Now i am not: The influence of pre-TMS oscillations and

corticomuscular coherence on motor-evoked potentials. *Cerebral Cortex*, *24*(7), 1708–1719. https://doi.org/10.1093/cercor/bht024

[45] Shadmehr, R., & Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, *14*(5 II), 3208–3224. https://doi.org/10.1523/jneurosci.14-05-03208.1994

[46] Shenoy, K. V., & Kao, J. C. (2021). Measurement, manipulation and modeling of brain-wide neural population dynamics. *Nature Communications*, *12*(1), 1–5. https://doi.org/10.1038/s41467-020-20371-1

[47] Shibata, K., Watanabe, T., Sasaki, Y., & Kawato, M. (2011). Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *Science*, *334*(6061), 1413–1415. https://doi.org/10.1126/science.1212003

[48] Soekadar, S. R., Witkowski, M., Birbaumer, N., & Cohen, L. G. (2015). Enhancing Hebbian Learning to Control Brain Oscillatory Activity. *Cerebral Cortex*, *25*(9), 2409–2415. https://doi.org/10.1093/cercor/bhu043

[49] Takemi, M., Masakado, Y., Liu, M., & Ushiba, J. (2013). Event-related desynchronization reflects downregulation of intracortical inhibition in human primary motor cortex. *Journal of Neurophysiology*, *110*(5), 1158–1166. https://doi.org/10.1152/jn.01092.2012

[50] Telgen, S., Parvin, D., & Diedrichsen, J. (2014). Mirror reversal and visual rotation are learned and consolidated via separate mechanisms: Recalibrating or learning de novo? *Journal of Neuroscience*, *34*(41), 13768–13779. https://doi.org/10.1523/JNEUROSCI.5306-13.2014

[51] Tomassini, A., Maris, E., Hilt, P., Fadiga, L., & D'Ausilio, A. (2020). Visual detection is locked to the internal dynamics of cortico-motor control. *PLOS Biology*, *18*(10), e3000898. https://doi.org/10.1371/journal.pbio.3000898

[52] Tsuchimoto, S., Shibusawa, S., Iwama, S., Hayashi, M., Okuyama, K., Mizuguchi, N., Kato, K., & Ushiba, J. (2021). Use of common average reference and large-Laplacian spatial-filters enhances EEG signal-to-noise ratios in intrinsic sensorimotor activity. *Journal of Neuroscience Methods*, *353*. https://doi.org/10.1016/j.jneumeth.2021.109089

[53] Van Der Maaten, L., Courville, A., Fergus, R., & Manning, C. (2014). Accelerating t-SNE using Tree-Based Algorithms. *Journal of Machine Learning Research*, *15*, 3221–3245. http://homepage.tudelft.nl/19j49/tsne;

[54] Van Der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579–2625.

[55] Wander, J. D., Blakely, T., Miller, K. J., Weaver, K. E., Johnson, L. A., Olson, J. D., Fetz, E. E., Rao, R. P. N., & Ojemann, J. G. (2013). Distributed cortical adaptation during learning of a brain-computer interface task. *Proceedings of the National Academy of Sciences*, *110*(26), 10818–10823. https://doi.org/10.1073/pnas.1221127110

[56] Witte, M., Kober, S. E., & Wood, G. (2018). Noisy but not placebo: Defining metrics for effects of neurofeedback. *Brain*, *141*(5), e40. https://doi.org/10.1093/brain/awy060

Figure Legends

Figure 1 Conceptual illustration of neural adaptation process induced by brain-computer interfacing

A: Setup of a brain-computer interface. Online acquired scalp electroencephalograms were fed into a classifier to detect the presence/absence of attempted movement. Predicted brain state was shown to participants as movement of visual object on display.

B: Conceptual visualization of cortical adaptation. Scaling adaptation reflects improvement in voluntary regulation of a specific component. If the centers of gravity determined from datapoints in two conditions are separated after brain-computer interfacing, it suggests the separability of two conditions is enhanced by adaptation. Deforming adaptation suggests that activity patterns are allocated to a specific brain state in order to adapt to the classifier. If the geometric relationships between two conditions are deformed with respect to a specific axis, it suggests the adaptation process progressed such that the two conditions are ´separated along the axis.

Figure 2 Temporal changes in acquired scores

Group results of performance scores from users of model-based (A), adaptive (B), and *de novo* (C) classifiers. Solid lines indicate mean values while shaded areas represent 1 standard deviation across participants.

Figure 2 Supplement 1

A: Electrode locations. The three classifiers used in the study had different channels of interest. The model-based classifier used only channel C3 indicated in blue around the left sensorimotor cortex. The adaptive classifier used whole-head EEG channels (purple) to construct a common spatial pattern. The *de novo* classifier used only the Cz channel, shown here in green.

B: Experimental protocol and time course of a trial

C: Visual feedback object. For the model-based or adaptive classifiers, an illustration of a hand was shown **t**hat matched the attempted movements of the users while an illustration of a tail was used in the *de novo* task to encourage users to acquire novel mental actions that enhanced controllability of the BCI.

Figure 2 Supplement 2

Results of source estimation analysis from representative participants. The colored regions indicate voxels where activities were significantly different during Rest and Imagine periods (p < 0.05 unc.). Areas colored with blue and green indicate those for model-based and de novo classifiers, respectively. While significant voxels were localized around the contralateral hemisphere of the imagined hand for the model-based classifier, those for the *de novo* classifier were located bilaterally, including in the pre/post central gyrus and supplementary motor area (peak voxel was in the postcentral gyrus, [MNI coordinates: -40, -25, 45]). Note that a representative source estimation for the adaptive classifier is not shown due to variable activity patterns among participants.

sLoreta analyses of statistical non-parametric mapping for estimated cortical sources of band power in the alpha band (8-13 Hz). Areas colored with blue and green indicate those from model-based and *de novo* classifiers, respectively. Masks superimposed on a standard brain template were visualized by MRIcroGL (https://www.mccauslandcenter.sc.edu/mricrogl/home).

Figure 3 low dimensional visualization of EEG data by t-SNE

A: Examples of t-SNE-based visualization of datasets from a representative participant in each classifier. Each axis represents results of the t-SNE analysis, which generates three axes from input data. Blue points represent data from the Imagine period and red ones are those from the Rest period.

B: Changes in geometric relationships between dataset and classifier plane. As training progressed, the geometric relationship of points from two brain states changed with respect to the classifier plane (black plane). The large points indicate the centers of gravity of points from each brain state. The black line orthogonal to the classifier plane is the classifier normal vector (see also Figure 3D)

C: An example of t-SNE-based data visualization in embedded space (Model-based classifier user). Each datapoint is colored with its SMR-ERD value derived from the C3 electrode around the left sensorimotor cortex. The black plane represents the classifier plane (see also equation 2.9 for mathematical details). The large points indicate the centers of gravity of points from each brain state. The black line orthogonal to the classifier plane is the classifier normal vector (see also Figure 3D).

D: The t-SNE-based quantification of the adaptation process with respect to the classifier plane. $tNorm_p$ is defined as a component of $tVec$ with respect to the

858 classifier vector, while θ_p is defined as a subtended angle between $tVec$ and
859 the classifier vector.
860
861 **Figure 4 Quantitative comparison of cortical adaptation processes in**
862 **embedded space**
863 Changes over time in $tNorm_p$ and θ_p for participants operating under the
864 model-based classifier (A), the *de novo* classifier (B), and the adaptive classifier
865 (C).
866
867 **Figure 4 Supplement 1** Quantitative comparisons of cortical adaptation
868 processes to the model-based classifier plane
869 Changes over time in $tNorm_p$ (A) and θ_p (B) of the model-based classifier for
870 participants operating under the adaptive (left, purple) or *de novo* (right, green)
871 classifiers.
872
873 **Figure 4 Supplement 2** Quantitative comparisons of cortical adaptation
874 processes to the de novo classifier plane.
875 Changes over time in $tNorm_p$ (left) and θ_p (right) of the *de novo* classifier for
876 participants operating under the model-based (A, blue) and adaptive (B, purple)
877 classifiers.
878
879 **Figure 4 Supplement 3** Quantitative comparisons of cortical adaptation
880 processes to the adaptive classifier plane.
881 Changes over time in $tNorm_p$ (left) *and* θ_p (right) of the adaptive classifier for
882 participants operating under the model-based (A, blue) and *de novo* (B, green)
883 classifiers.
884

Supplementary Material for: *De novo* brain-computer interfacing deforms manifold of population activity patterns of human cerebral cortex

Detailed procedure for construction of adaptive classifier

In the experiments with the adaptive classifier, six spatial filters were generated via the common spatial pattern algorithm during the intervals between sessions. Filters that maximized variance differences were generated via the CSP algorithm and applied to the online EEG signals during the subsequent session. The log-transformed variances of the six-channel, spatial-filtered data from the previous 1-s signals were calculated and classified with a linear SVM classifier.

Cross validation results of the adaptive decoder are shown in Figure S1A. Data from a single session were used for training and data from the remaining sessions were used for testing. Performance, evaluated from the coefficients of linear regression, did not show systematic improvement at the group level (p = 0.078, Wilcoxon sign-rank test). Figure S1 B shows the results of a cross validation test using a single session for training and another for testing, suggesting there was an increase in accuracy during the later period, which was confirmed by temporal changes in the acquired score (Figure 2B).

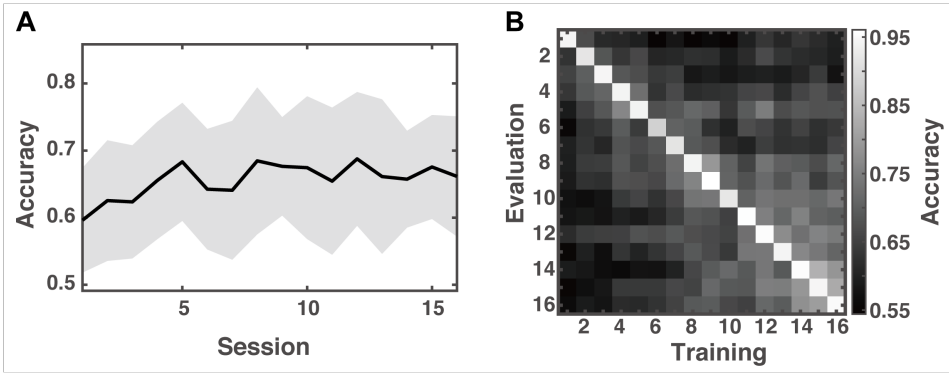


Figure S1. Temporal changes in cross validation performance of adaptive decoder

Cortical source estimation analysis for EEG signals

For the data from participants who operated under the model-based or *de novo* decoders, band-power in the alpha-band was subjected to sLORETA analysis for cortical source estimation (Pascual-Marqui, 2002). Because the motivation for conducting the source analysis was to test whether the targeted region of the

classifier was successfully activated during the late period of BCI training, averaged data from across the last four sessions were subjected to a non-parametric permutation test (Nichols & Holmes, 2002).

Mathematical description of t-SNE algorithm

In the original manuscript describing the t-SNE algorithm (Van Der Maaten & Hinton, 2008), the datapoints are described as $\boldsymbol{x} = (x_1, x_2, \ldots x_n)$, where x_i is a vector with the arbitrary number of features. Assume that x_i and x_j are two data points, d_{ij} is the distance between the two points and $p_{j|i}$ is the probability that x_i. and x_j are neighbors. The probabilities follow a Gaussian distribution, described by:

$$\boldsymbol{p_{j|i}} = \frac{\mathbf{exp}(-\boldsymbol{d_{ij}^2}/\mathbf{2}\boldsymbol{\sigma_i^2})}{\boldsymbol{\Sigma_{k\neq i}}\,\mathbf{exp}(-\boldsymbol{d_{ik}^2}/\mathbf{2}\boldsymbol{\sigma_i^2})}, \qquad \boldsymbol{p_{i|i}} = \mathbf{0}$$

where σ_i is determined by the parameter *perplexity*, the value of which was calculated with H:

$$\mathrm{perplexity}(P_i) = 2^{H(P_i)}$$

$$\boldsymbol{H(P_i)} = -\sum_{\boldsymbol{j}} \boldsymbol{p_{j|i}}\,\mathbf{log_2}\,\boldsymbol{p_{j|i}}$$

The value of σ_i is adjusted in a binary search method so that *perplexity* matches a value determined by the user. According to Van Der Maaten & Hinton (2008), perplexity is a smooth measure of the effective number of neighbors. The joint probability p_{ij} is defined by symmetrizing the conditional probabilities:

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N},$$

where N is the number of data points. The pairwise distances between points in low-dimensional embedding were converted to possibilities that follow a Student's t-distribution with one degree of freedom. Assume that y_i and y_j are two data points in the embedded space and joint possibilities q_{ij} are defined as:

$$q_{ij} = \frac{\left(1 + d_{ij}^2\right)^{-1}}{\Sigma_{k \neq l}\left(1 + d_{kl}^2\right)^{-1}}$$

The Kullback-Liebler (KL) divergence between joint possibility distribution *P* in the original space and *Q* in embedded space was then calculated with:

$$KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}$$

The KL divergence was minimized via a gradient descent method by adjusting the positions of points in the embedded space.

References

[1] Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, *15*(1), 1–25. https://doi.org/10.1002/hbm.1058

[2] Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, *24*(SUPPL. D), 5–12.

[3] Van Der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579–2625.