Perceptual Learning Incepted by Decoded fMRI Neurofeedback Without Stimulus Presentation

Kazuhisa Shibata,* Takeo Watanabe,*† Yuka Sasaki,‡ Mitsuo Kawato

It is controversial whether the adult primate early visual cortex is sufficiently plastic to cause visual perceptual learning (VPL). The controversy occurs partially because most VPL studies have examined correlations between behavioral and neural activity changes rather than cause-and-effect relationships. With an online-feedback method that uses decoded functional magnetic resonance imaging (fMRI) signals, we induced activity patterns only in early visual cortex corresponding to an orientation without stimulus presentation or participants' awareness of what was to be learned. The induced activation caused VPL specific to the orientation. These results suggest that early visual areas are so plastic that mere inductions of activity patterns are sufficient to cause VPL. This technique can induce plasticity in a highly selective manner, potentially leading to powerful training and rehabilitative protocols.

Whether adult primate visual cortex has sufficient plasticity to allow for behavioral and/or sensitivity changes remains a point of great controversy. Many studies have examined how activity changes in the brain are correlated with performance improvements on a visual task resulting from repetitive training, known as visual perceptual learning (VPL). However, such a correlational approach has not conclusively settled the adult plasticity debate. Although some studies have found correlations between performance increase and changes in early visual areas (*1*–*5*), other studies found correlations in higher visual and/or decision areas (*6*–*8*). None of these studies directly addresses the question of whether early visual areas are sufficiently plastic to cause VPL. Changes in early visual areas observed in correlation with VPL do not exclude the possibility that the changes are, in reality, a reflection of the influences of changes in other brain areas. On the other hand, changes in higher brain areas in some conditions in correlation with some types of VPL do not rule out the possibility that early areas are sufficiently plastic to cause VPL in other conditions.

To addresses the question of whether early visual areas are that plastic, we developed a functional magnetic resonance imaging (fMRI) online-feedback method, by which activation patterns corresponding to the pattern evoked by the presentation of a real and specific target orientation stimulus were repeatedly induced without the participants' knowledge of what is being learned and without external stimulus presentation [see supporting online materials (SOM) and methods]. The mere induction of the activation patterns resulted in significant behavioral performance improvement on the target stimulus orientation, but not on other orientations.

The complete experiment consisted of four stages: (i) pre-test (1 day), (ii) fMRI decoder construction (1 day), (iii) induction (decoded fMRI neurofeedback, 5 days for four participants and 10 days for six participants), and (iv) post-test (1 day) stages (Fig. 1A).

In the pre- and post-test stages, we measured participants' performance in an orientation discrimination task (10 participants). In each trial (Fig. 1B), participants were asked to report which of three orientations (10°, 70°, or 130°) (Fig. 1C) had been presented in a Gabor patch (see SOM).

Next, we tried to obtain fMRI activity patterns in V1/V2 that are induced by the presentation of each of the three tested orientations in Gabor patches for each participant (fMRI decoder construction stage). Participants were asked to perform a task designed to maintain their attention to the Gabor patches while fMRI signals in V1/V2 were measured (Fig. 1D and SOM). Based on the fMRI signals, we constructed a multinomial sparse logistic regression decoder (*9*) to classify a pattern of the measured fMRI signals into one of the three orientations (fig. S1).

Once the decoder was constructed, each participant took part in a 5- or 10-day induction stage during which he or she learned to induce activation patterns in V1/V2 that corresponded to the target orientation, one of the three orientations, that was randomly assigned to each participant. During each trial (Fig. 1E and SOM), participants were asked to "somehow regulate activity in the posterior part of the brain to make the solid green disc that was presented 6 s later as large as possible (the maximum possible size corresponds to the outer green circle)." The size of the disc presented in the feedback period corresponded to the decoder output for the target orientation, which represented the likelihood of the blood oxygen level–dependent signal pattern in V1/V2 obtained in the preceding induction period being classified into the target orientation for which the performance is aimed to be improved. It roughly represented how similar the activation pattern obtained in the induction period in the absence of visual stimulation is to the pattern evoked by the real Gabor stimulus of the

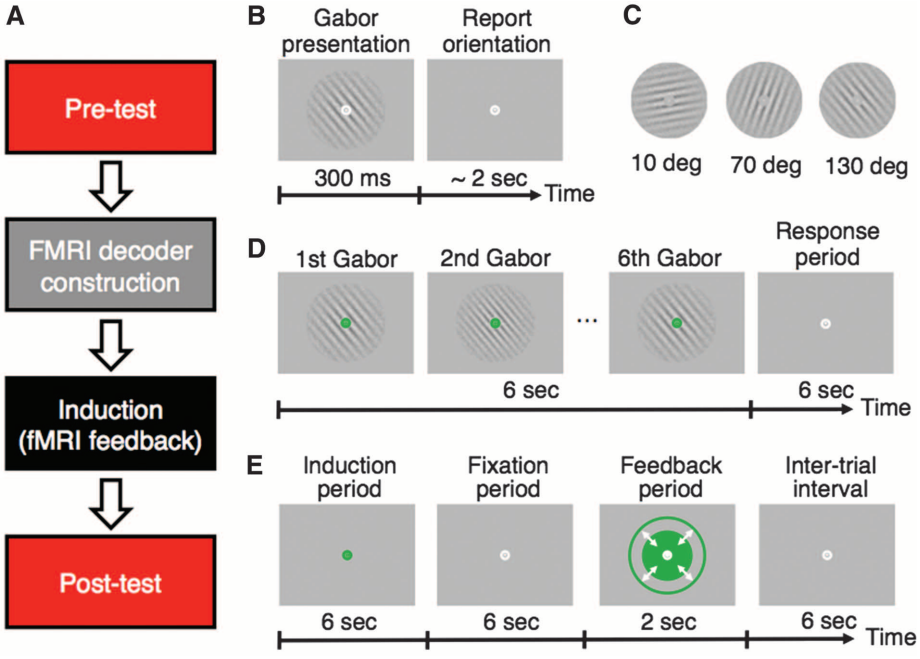

Fig. 1. Procedure of the main experiment. (**A**) Stages of the experiment. (**B**) Procedure of a trial of the test stages. (**C**) Gabor patches of the three orientations used in this experiment. (**D**) Example of a sequence of stimuli in one trial of the decoder construction stage. (**E**) Procedure of a trial in the induction stage.

Advanced Telecommunications Research Institute International Computational Neuroscience Laboratories, 2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan.

*Present address: Department of Psychology, Boston University, 64 Cummington Street, Boston, MA 02215, USA.
†To whom correspondence should be addressed. E-mail: takeo@bu.edu
‡Present address: Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, 149 Thirteenth Street, Charlestown, MA 02129, USA; and Department of Radiology, Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, USA.

target orientation presented during the decoder construction stage. We call the likelihood (similarity) target-orientation likelihood. However, participants were not informed of what the size represented. They were told that they would receive a payment bonus proportional to the mean size of the feedback disc, but all other information, including the target orientation, the purpose of the neurofeedback, and the meaning of the disc size, was withheld from the participants.

The main purpose of our study was to test whether early visual cortical areas are sufficiently plastic to cause VPL of a specific orientation as a result of mere repetitive inductions of activity patterns corresponding to that orientation. However, before testing this hypothesis, it was necessary to examine whether participants could learn to induce significantly high target-orientation likelihood; that is, a neural activity pattern in V1/V2 that is similar to an activity pattern evoked by the actual presentation of the target orientation. Thus, we first examined whether outputs of the decoder could be biased toward the selected target orientation compared with the other two orientations that were rotated ±60° from the target orientation. Figure 2 shows that the overall mean target-orientation likelihood in V1/V2 was significantly higher than chance across the participants, on average, during the induction stage [$t(9) = 3.34$, $P < 10^{-2}$]. The mean (across the participants) target-orientation likelihood in V1/V2 for the first 30 trials of the first neurofeedback day was around chance level (fig. S2). Thus, there was no significant orientation bias for the target orientation before neurofeedback, and participants quickly learned to induce significantly high target-orientation likelihood, even during the first neurofeedback day (see performance for day 1 in Fig. 2). We also applied the same analysis to the overall activity pattern that is the mean across trials and confirmed the same tendency (fig. S3).

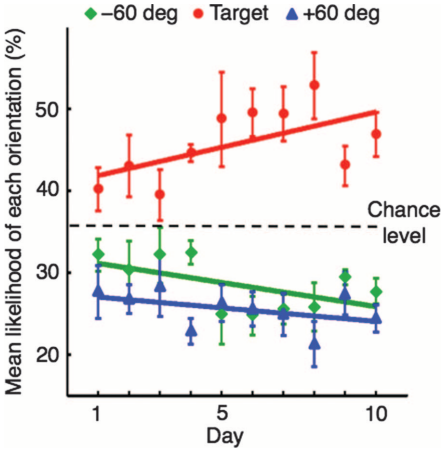

Fig. 2. Results of the induction stage. The mean (±SE; represented by error bars) likelihoods of three orientations assessed by a decoder based on V1/V2 activity patterns. For the first 5 days, combined data from 5-day (4 participants) and 10-day (6 participants) neurofeedback sessions is shown. For the last 5 days, data from 10-day neurofeedback session is shown.

Were participants aware of the purpose of the induction stage? After the post-test stage, participants were asked about what they thought the size of the feedback disc represented, but none of their responses was even remotely related to the true workings of the experiment (see "After post-test stage" section of SOM). After being told that the disc size represented the likelihood of one of three orientations, subjects were asked to report the orientation they thought they had been trained on by picking one of the three orientations. The percentage of the choice of the target orientation in fig. S4 was statistically undistinguishable from what would be expected from chance ($\chi^2 = 0.20$, $P = 0.90$).

The purpose of the induction stage was to have participants learn and then continue to induce activity patterns with significantly high target-orientation likelihood in V1/V2. This learning in the induction stage should not be confused with the plasticity or VPL that refers to improvements on visual tasks. The main purpose of the present study was to examine whether the mere repetitive induction of specific activation patterns in V1/V2 causes VPL reflected as performance improvement.

We compared participants' performance in the pre- and post-test stages (Fig. 3, A to C). Three-way [test stage × orientation × signal-to-noise (S/N) ratio] analysis of variance with repeated measures indicated significant main effect of S/N ratio ($F_{3,27} = 683.17$, $P < 10^{-4}$) and significant effect of interaction between test stage, orientation, and S/N ratio ($F_{6,54} = 2.68$, $P = 0.02$). Post-hoc t test between accuracies in pre- and post-tests revealed that discrimination performance for the target orientation significantly improved at the 6% S/N ratio [$t(9) = 5.76$, $P < 10^{-2}$ with Bonferroni correction]. Sensitivity (d′) in the pre-test subtracted from that of the post-test was significantly greater than zero for the target orientation at the 6% S/N ratio [$t(9) = 5.60$, $P < 10^{-3}$ with Bonferroni correction] (Fig. 3D).

What is the relation between the target-orientation likelihood in V1/V2 and d′ changes? The sensitivity changes for the participants with 10 days training (induction) were larger than for those with 5 days training (fig. S5). This observation was consistent with the general tendency that the magnitude of VPL is larger with longer training until it reaches an asymptote. Thus, we computed the summation of the target-orientation likelihoods in V1/V2 for all trials for each participant and plotted the sensitivity change against the summation. The correlation was even stronger for the likelihood summation ($r = 0.87$, $P = 10^{-3}$) (Fig. 3E) than for the average likelihood ($r = 0.74$, $P = 0.01$).

To test whether the VPL observed in the main experiment resulted simply from participants' involvement in the test stages, we conducted a control experiment with six new participants in which only the pre- and post-test stages were conducted (SOM). No significant performance improvement was observed (fig. S6).

Participants can be trained to control the overall mean activation of an entire brain region or the activation in one region relative to that in another region (*10*–*13*). One might wonder if the participants in the present study simply learned to regulate the overall activity of V1/V2. However, two lines of evidence argue against this possibility. First, the multinomial sparse logistic regression decoder used in this study computed the linear weighted sum of voxel activities, and weights of the decoder were almost symmetrically distributed around 0 (fig. S7). Second, some voxels in V1/V2 were activated positively and others negatively, rather than uniform positive or negative activation during the induction stage (fig. S8).

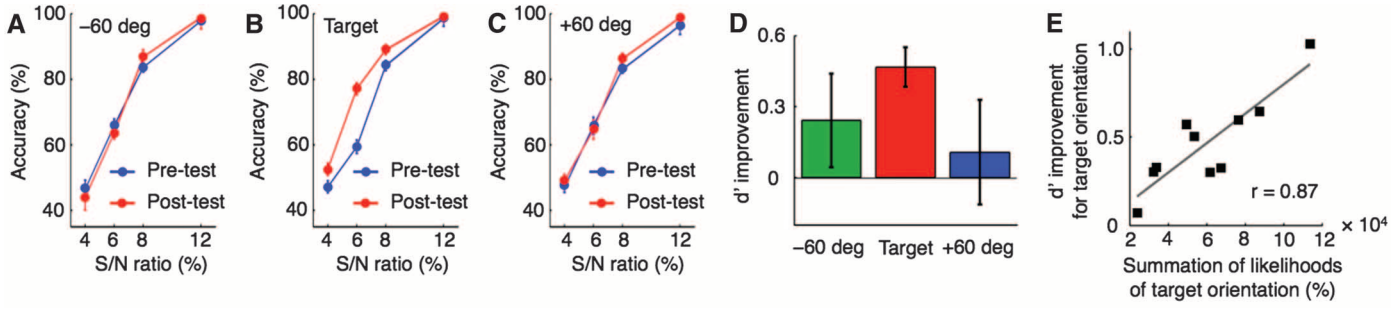

Fig. 3. Results of the pre- and post-tests. (**A** to **C**) Discrimination accuracies for –60° (A), target orientation (B), and +60° (C). Error bars indicate SE. (**D**) Sensitivity values (d′) in the pre-test subtracted from those in the post-test for three orientations at 6% S/N ratio. Error bars indicate SE. (**E**) Correlation between the summation of target-orientation likelihoods in V1/V2 in the induction stage and sensitivity changes for the target orientation at 6% S/N ratio.

The neurofeedback provided to participants was based on activation patterns only in V1/V2. However, this procedure might have induced neural activities in areas other than V1/V2, which might also contribute to VPL. To test whether other regions quantitatively contributed to VPL, we conducted two offline tests with other areas (such as V3, V4, the intraparietal sulcus, and the lateral prefrontal cortex) that have been implicated in VPL (*6*–*8*).

If the orientation-specific activation patterns in V1/V2 during the induction stage induced similar orientation-specific brain activities in other areas, the activation patterns in those areas should predict the target-orientation likelihood in V1/V2 on a trial-by-trial basis. In the first offline test, we employed a sparse linear regression method (*14*) to predict the target-orientation likelihoods in V1/V2 from activation patterns in those higher areas in each trial during the induction stage (SOM). Goodness of prediction for the target-orientation likelihood in V1/V2 by other areas, or prediction accuracies of the sparse linear regression, was evaluated by coefficients of determination, all of which were less than 5% (fig. S9A).

We conducted a second offline test to examine the possibility that the decoder simply performed poorly in higher brain areas. We examined whether accurate orientation information can be read out from each brain area when real orientation stimuli are presented in the decoder construction stage. As was done for V1/V2 during the fMRI decoder construction stage, we built a multinomial sparse logistic regression decoder to classify activation patterns into each of the three orientations (SOM). Decoding accuracies were significantly higher than chance level in all of these areas (fig. S9B, also compare fig. S9, A and B). The results of these two offline tests indicate that influences of the neurofeedback on VPL were largely confined to early visual areas such as V1/V2.

Our results indicate that the adult early visual cortex is so plastic that mere repetition of the activity pattern corresponding to a specific feature in the cortex is sufficient to cause VPL of a specific orientation, even without stimulus presentation, conscious awareness of the meaning of the neural patterns that participants induced, or knowledge of the intention of the experiment. How is the present research on VPL distinguished from previous approaches? Unit recording and brain imaging studies have successfully revealed the correlation between VPL and neural activity changes (*1*–*8*). However, these correlation studies cannot clarify cause-and-effect relationships. The studies that examined the effect of a lesion (*15*) or transcranial magnetic stimulation (TMS) (*16*, *17*) to a brain region on VPL have shown whether the examined region plays some role in VPL. However, these studies cannot clarify how particular activity patterns in the region are related to VPL. In contrast, the present decoded fMRI neurofeedback method allowed us to induce specific neural activity patterns in V1/V2, which caused VPL.

The present decoded fMRI neurofeedback method can be used to clarify cause-and-effect relationships in many functions in system neuroscience (*18*, *19*). Although previous fMRI online-feedback training is a promising technique for influencing human behaviors (*10*–*13*), as in lesion or TMS studies, it could at best reveal influences of the entire extent of an area/region on learning/memory, which is a certain limitation for neuroscientific research (*20*). In contrast, the present decoded fMRI neurofeedback method induces highly selective activity patterns within a brain region, thus allowing the investigator to influence specific functions. It can "incept" a person to acquire new learning, skills, or memory, or possibly to restore skills or knowledge that has been damaged through accident, disease, or aging, without a person's awareness of what is learned or memorized.

References and Notes

1. A. Schoups, R. Vogels, N. Qian, G. Orban, *Nature* **412**, 549 (2001).
2. Y. Yotsumoto, T. Watanabe, Y. Sasaki, *Neuron* **57**, 827 (2008).
3. T. Hua *et al.*, *Curr. Biol.* **20**, 887 (2010).
4. N. Censor, Y. Bonneh, A. Arieli, D. Sagi, *J. Vision* **9**, 1 (2009).
5. A. Karni, D. Sagi, *Nature* **365**, 250 (1993).
6. C. T. Law, J. I. Gold, *Nat. Neurosci.* **11**, 505 (2008).
7. T. Yang, J. H. Maunsell, *J. Neurosci.* **24**, 1617 (2004).
8. C. M. Lewis, A. Baldassarre, G. Committeri, G. L. Romani, M. Corbetta, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 17558 (2009).
9. O. Yamashita, M. A. Sato, T. Yoshioka, F. Tong, Y. Kamitani, *Neuroimage* **42**, 1414 (2008).
10. S. Bray, S. Shimojo, J. P. O'Doherty, *J. Neurosci.* **27**, 7498 (2007).
11. A. Caria *et al.*, *Neuroimage* **35**, 1238 (2007).
12. R. C. deCharms *et al.*, *Neuroimage* **21**, 436 (2004).
13. N. Weiskopf *et al.*, *Neuroimage* **19**, 577 (2003).
14. A. Toda, H. Imamizu, M. Kawato, M. A. Sato, *Neuroimage* **54**, 892 (2011).
15. K. R. Huxlin *et al.*, *J. Neurosci.* **29**, 3981 (2009).
16. E. Corthout, B. Uttl, V. Walsh, M. Hallet, A. Cowey, *Neuroreport* **11**, 1565 (2000).
17. F. Giovannelli *et al.*, *Neuropsychologia* **48**, 1807 (2010).
18. H. R. Dinse, P. Ragert, B. Pleger, P. Schwenkreis, M. Tegenthoff, *Science* **301**, 91 (2003).
19. Y. Miyashita, *Science* **306**, 435 (2004).
20. M. Kawato, *Philos. Trans. R. Soc. London Ser. B* **363**, 2201 (2008).

Acknowledgments: This work was conducted in "Brain Machine Interface Development" under the Strategic Research Program for Brain Sciences by the Ministry of Education, Culture, Sports, Science and Technology of Japan. T.W. was partially supported by NIH grants R01 AG031941 and R01 EY015980 and Y.S. by grants R01 MH091801 and NSF 0964776. We thank J. Dobres, M. Fukuda, G. Ganesh, H. Imamizu, A. R. Seitz, and K. Tanaka for their comments on a draft of this manuscript and M. Fukuda, Y. Furukawa, S. Hirose, M. Sato, and ATR BAIC for technical assistances.

Supporting Online Material

www.sciencemag.org/cgi/content/full/334/6061/1413/DC1
Materials and Methods
Figs. S1 to S9
References (*21*–*31*)

1 August 2011; accepted 26 October 2011
10.1126/science.1212003

Entorhinal Cortex Layer III Input to the Hippocampus Is Crucial for Temporal Association Memory

Junghyup Suh,[1] Alexander J. Rivest,[1] Toshiaki Nakashiba,[1] Takashi Tominaga,[2] Susumu Tonegawa[1]*

Associating temporally discontinuous elements is crucial for the formation of episodic and working memories that depend on the hippocampal-entorhinal network. However, the neural circuits subserving these associations have remained unknown. The layer III inputs of the entorhinal cortex to the hippocampus may contribute to this process. To test this hypothesis, we generated a transgenic mouse in which these inputs are specifically inhibited. The mutant mice displayed significant impairments in spatial working-memory tasks and in the encoding phase of trace fear-conditioning. These results indicate a critical role of the entorhinal cortex layer III inputs to the hippocampus in temporal association memory.

A critical feature of episodic memory shared by some forms of working memory is the ability to associate temporally discontinuous elements, called temporal association memory (*1*–*3*). However, the neural circuits within the entorhinal cortex (EC)–hippocampus (HP) network subserving this type of association have remained unknown. The EC provides inputs to the HP via two major projections (Fig. 1A): the trisynaptic pathway (TSP) originating from EC layer II and the monosynaptic pathway (MSP) originating from EC layer III (ECIII). Studies on genetically engineered mice (*4*–*7*) and lesioned rats (*8*–*11*) have demonstrated crucial roles of the TSP in several features of episodic-memory processing, such as pattern completion (*5*–*8*) and separation (*4*, *8*). In contrast, the MSP contributions to episodic-memory processing remain poorly known. We tested the

[1]The Picower Institute for Learning and Memory, RIKEN–MIT Center for Neural Circuit Genetics, Department of Biology and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]Department of Neurophysiology, Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Kagawa, Japan.

*To whom correspondence should be addressed. E-mail: tonegawa@mit.edu

Supporting Online Material for

Perceptual Learning Incepted by Decoded fMRI Neurofeedback Without Stimulus Presentation

Kazuhisa Shibata, Takeo Watanabe,* Yuka Sasaki, Mitsuo Kawato

*To whom correspondence should be addressed. E-mail: takeo@bu.edu

Published 9 December 2011, *Science* **334**, 1413 (2011)
DOI: 10.1126/science.1212003

This PDF file includes:

Materials and Methods

The complete experiment consisted of 4 stages: pre-test (1 day), fMRI decoder construction (1 day), induction (decoded fMRI neurofeedback, 5 days for 4 subjects and 10 days for 6 subjects), and post-test (1 day) stages (Fig. 1A). Different stages were separated by at least 24 hours.

Subjects

Sixteen naïve subjects (20 to 38 years old; 11 males and 5 females) with normal or corrected-to-normal vision participated in the study. All the experiments and data analyses were conducted in Advanced Telecommunications Research Institute International (ATR). The study was approved by the Institutional Review Board of ATR. All subjects gave written informed consents.

Pre- and post-test stages

Only behavioral data were collected in the pre- and post-test stages. Oriented Gabor patches (spatial frequency=1 cycle/deg, contrast=100%, sigma of its Gaussian filter=2.5 deg, random spatial phase) were presented within an annulus subtending 0.75 to 5 deg from the center of a gray screen and were spatially masked by noise, which was generated from a sinusoidal luminance distribution, at a certain signal to noise (S/N) ratio (*21*). For example, in the case of 12% S/N ratio, 88% of pixels of a Gabor patch were replaced by noise.

Subjects' discrimination performance for the Gabor patches was measured (Fig. 1B) before (pre-test stage) and after (post-test stage) the induction stage (see below). In each trial, one of three Gabor orientations (10, 70, 130 deg; Fig. 1C for examples) was presented at one of four S/N ratios (4, 6, 8, 12%). The order of presentations of the orientations and S/N ratios was randomly determined and counterbalanced across trials. Throughout the task, the subjects were asked to fixate their eyes on a white bull's-eye on a gray disc (0.75 deg radius) at the center of the display. Each trial started with a 750-ms fixation period. Then, a Gabor patch was presented for 300 ms, after which the subjects were given 2 sec to report which of 3 possible Gabor orientations was presented by pressing one of 3 buttons on a keyboard. After each trial, a 500-ms inter-trial interval was inserted, consisting of a blank gray background. A brief break period was provided after each run of 50 trials. The subjects conducted 600 trials on each day. The magnitude of performance improvement was defined as the performance in the pre-test subtracted from the performance in the post-test.

FMRI decoder construction stage

The purpose of the fMRI decoder construction stage was to obtain the fMRI signals corresponding to each of three actual orientations, which would then be used to compute the parameters for the decoder used in the later induction stage.

First, we measured subjects' retinotopic maps to delineate visual cortical areas individually using a standard retinotopic method with blood-oxygen-level dependent (BOLD) signal (see elsewhere (*22-24*) for more details; see *MRI parameters* below). In addition, the subjects were presented with a reference stimulus to localize the retinotopic regions in V1/V2 corresponding to the visual field stimulated by the Gabor patches. The

reference stimulus was composed of a colored checkerboard pattern presented within an annulus subtending 1 to 4.75 deg from the center of a gray screen. We used a smaller annular region for the reference stimulus than for the Gabor patches to avoid selecting voxels corresponding to the stimulus edges, which may contain information irrelevant to orientation (*25*).

Next, we measured subjects' BOLD signal patterns (see *MRI parameters* below) for the three Gabor orientations (10, 70, 130 deg) used in the pre- and post-test stages at 50% S/N ratio. Throughout the fMRI run, the subjects were asked to fixate their eyes on a white bull's-eye on a gray disc (0.75 deg radius) presented at the center of the display. The subjects conducted 240 trials in a total of ten fMRI runs. A brief break period was provided after each run upon subjects' requests.

Each fMRI run consisted of 24 task trials (1 trial=12 sec; Fig. 1D), plus a 10-sec fixation period before the trials and 2-sec fixation after the trials (1 run=300 sec). The fMRI data for the initial 10 sec of each run were discarded due to possible unsaturated T1 effects. Each task trial consisted of a stimulus period (6 sec) and a response period (6 sec). At the beginning of each task trial, the color of the fixation point changed from white to green to indicate the start of the stimulus period, during which Gabor patches with one of the three orientations flashed at 1 Hz. Thus, during the stimulus period, the same Gabor orientation was presented 6 times. One of the three Gabor orientations was randomly assigned to each trial. In a half of 24 trials of each fMRI run, the spatial frequency of one of the flashing Gabor patches was slightly increased relative to the other 5 Gabor patches. In the other half, the spatial frequency of the flashing Gabor patches did not change.

The stimulus period was followed by the response period, in which only the fixation point was presented. The color of the fixation point reverted to white. During the response period, the subjects were asked to respond as to whether there was any spatial frequency change in the Gabor patches presented in the prior stimulus period. The subjects were instructed to press the button with their right hand if they detected any spatial frequency change (for example, the second Gabor in Fig. 1D consists of a higher spatial frequency). If subjects did not detect any frequency change, they were instructed not to press the button.

Task difficulty was controlled according to subjects' button responses using an adaptive staircase method (*26*) so that the task difficulty was kept constant throughout trials: the degree of spatial frequency change for the next trial was decreased by 0.02 Hz in the case of a correct detection (hit), and increased by 0.02 Hz in the case of a false alarm or miss. In the case of a correct rejection, the degree of spatial frequency change remained the same. The mean (±s.e.) degree of the spatial frequency change during the task was 0.1174±0.009 Hz, and the mean (±s.e.) task accuracy was 71.3±0.7% across the subjects.

Measured fMRI signals to Gabor orientations and to stimuli for retinotopy were preprocessed using BrainVoyager QX software. All functional images underwent 3D motion correction. No spatial or temporal smoothing was applied. Rigid-body transformations were performed to align the functional images to the structural image for each subject. A gray matter mask was used to extract fMRI data only from gray matter voxels for further analyses.

We conducted a voxel-by-voxel conventional amplitude analysis (*23, 24*) to identify the retinotopic region of V1, V2, and the sub-region that corresponded to the reference stimulus within V1/V2 (the reference region). Once we identified the reference region, time-courses of BOLD signal intensities were extracted from each voxel in the reference region and shifted by 6 sec to account for the hemodynamic delay using the Matlab software. A linear trend was removed from the time-course, and the time-course was z-score normalized for each voxel in each run to minimize baseline differences across the runs. The data samples for computing the decoder were created by averaging the BOLD signal intensities of each voxel for 3 volumes corresponding to the 6-sec stimulus period.

We used a multinomial linear sparse logistic regression, which automatically selected the relevant voxels in the reference region within V1/V2 for decoding (*27, 28*), to construct a decoder for the induction stage. The input voxels were selected from the region of the 1.00 to 4.75 deg eccentricity within V1/V2. Thus, the voxels at the fovea and those that reflected stimulus edges were excluded from the analyses. Because of this exclusion and the usage of the Gabor stimulus with 5 deg radius, T-junctions were doubly avoided as a source of orientation information. We trained the decoder to classify a pattern of BOLD signals into one of three Gabor orientations (10, 70, 130 deg) using 240 data samples obtained from 240 trials in the ten fMRI runs. As a result, the inputs to the decoder were the subjects' moment-to-moment brain activations, while the outputs of the decoder represented the calculated likelihood of each Gabor orientation being presented to the subjects. The mean (±s.e.) number of voxels for decoding was 239±29 in V1/V2.

Induction stage

In the induction stage, which consisted of 5 or 10 daily sessions (but not necessarily consecutive), the subjects were instructed to regulate activation in the posterior part of the brain without any actual visual stimuli presented except for the central fixation point. Debriefing interviews conducted after the experiment confirmed that the subjects were naïve about the function of the posterior part of the brain.

During each neurofeedback day, subjects participated in up to 12 fMRI runs. The mean (±s.e.) number of runs in each day was 10.8±0.3 across days and subjects. Each fMRI run consisted of 15 trials (1 trial=20 sec) preceded by a 30-sec fixation period (1 run=330 sec). The fMRI data for the initial 10 sec were discarded to avoid unsaturated T1 effects. Throughout a run, the subjects were instructed to fixate their eyes on a white bull's-eye at the center of a gray disc (0.75 deg radius) presented at the center of the display. After each run, a brief break period was provided upon a subject's request.

Each trial (Fig. 1E) consisted of an induction period (6 sec), a fixation period (6 sec), a feedback period (2 sec), and an inter-trial interval (6 sec) in this order. During the induction period, the color of the fixation point changed from white to green. No visual stimulus except the fixation point was presented during the induction period. The subjects were instructed to regulate the posterior part of their brains, with the goal of making the size of a solid green disc presented in the later feedback period as large as possible. The experimenters provided no further instructions or strategies. During the fixation period, the subjects were asked simply to fixate on the central point. This period was inserted between the induction period and the feedback period due to the known hemodynamic delay, which we assumed lasted 6 sec, during which V1/V2 activation patterns were

calculated in time for a green disc to be shown in the subsequent feedback period. During the feedback period, the green disc was presented for 2 sec. The size of the disc presented in the feedback period represented how much a momentary BOLD signal pattern in V1/V2 obtained in the prior induction period corresponded to the pattern induced by the presentation of the real and specific targeted Gabor orientation, collected through the above-mentioned fMRI decoder construction stage. The green disc was always enclosed by a larger green concentric circle (5 deg radius), which indicated the disc's maximum possible size. The feedback period was followed by an inter-trial interval that lasted 6 sec, during which the subjects were asked to fixate on a central white point. This period was followed by the start of the next trial.

The target Gabor orientation was randomly selected from one of the three orientations (10, 70, 130 deg) and assigned to each subject without informing the subjects about the assigned target orientation. The remaining two orientations were rotated from -60 deg and +60 deg from the target orientation. The size of the disc presented during the feedback period was computed during the fixation period in the following manner. First, measured functional images during the induction period underwent 3D motion correction using Turbo BrainVoyager. Second, time-courses of BOLD signal intensities were extracted from each of the voxels identified in the fMRI decoder construction stage, and were shifted by 6 sec to account for the hemodynamic delay. Third, a linear trend was removed from the time-course, and the BOLD signal time-course was z-score normalized for each voxel using BOLD signal intensities measured for 20 sec starting from 10 sec after the onset of each fMRI run. Fourth, the data sample to calculate the size of the disc was created by averaging the BOLD signal intensities of each voxel for 6 sec in the induction period. Finally, the likelihood of each orientation was calculated from the data sample using the decoder computed in the fMRI decoder construction stage. The size of the disc was proportional to the likelihood (ranging from 0 to 100%) of the target orientation assigned to each subject. The target orientation was constant throughout the induction stage.

In addition to a fixed amount of the compensation for participation in the experiment, a bonus of up to $30 was paid to the subjects based on the mean size of the disc in each day.

After post-test stage

When the subjects were asked what they tried to do to increase the size of the feedback disc, the most common reply was: "I tried various things since I had no idea of the correct way". The strategies each subject reported to use were as follows:

Subject 1 " I tried to remember various scenes from one famous animated movie."
Subject 2 "I tried to remember various things that had happened yesterday, or to image I am moving my fingers."
Subject 3 "I tried to keep switching my attention between the two eyes or to imagine a big green disc."
Subject 4 "I tried to focus my attention on the fixation point at the center of the display."
Subject 5 "I tried to obtain and keep an image of a big green disc."
Subject 6 "I tried to keep my attention on the back part of my brain."
Subject 7 "I tried to focus my attention on the color of the fixation point at the center of the display."

Subject 8 “I tried to imagine a big green disc.”
Subject 9 “I tried to imagine a big green disc.”
Subject 10 “I tried to remember old memories or to perform numerical calculation.”

Apparatus and stimuli

Visual stimuli were presented on a LCD display (1024 × 768 resolution, 60 Hz refresh rate) during the pre- and post-test stages and via a LCD projector (1024 × 768 resolution, 60 Hz refresh rate) during fMRI measurements in a dim room. All visual stimuli were made using Matlab and Psychtoolbox 3 (*29*) on Mac OS X.

MRI parameters

The subjects were scanned in a 3T MR scanner with a head coil in the ATR Brain Activation Imaging Center. Retinotopy, fMRI signals for the fMRI decoder construction and induction stages were acquired using a gradient EPI sequence. In all fMRI experiments, 33 contiguous slices (TR=2 sec, voxel size=3 × 3 × 3.5 mm^3, 0 mm slice gap) oriented parallel to the AC-PC plane were acquired, covering the entire brain. For an inflated format of the cortex used for retinotopic mapping and an automated parcellation method (*30*), T1-weighted MR images (MP-RAGE; 256 slices, voxel size=1 × 1 × 1 mm^3, 0 mm slice gap) were also acquired during the decoder construction stage.

Control experiment

The purpose of the control experiment was to test whether participating in the pre- and post-test stages alone was sufficient to improve subjects’ discrimination performance, without an induction stage. Thus, the control experiment consisted only of the pre- and post-test stages. The procedure was otherwise identical to that of the main experiment. Since the average interval between the pre- and post-test stages in the main experiment with 10-day induction stage was 28 days, the interval between the pre- and post-test stages in the control experiment was also set to this length.

Offline tests

We conducted two offline tests for the following 5 additional brain areas: V3, V4, the intraparietal sulcus (IPS), the lateral prefrontal cortex (LPFC), and the combined region from V3 to LPFC. Localizations of V3 and V4 were determined based on individual retinotopic maps. Localizations of IPS and LPFC were determined based on anatomical landmarks derived from an automated brain parcellation method (*30*). LPFC was anatomically defined as the middle frontal gyrus plus the inferior frontal sulcus.

In the first offline test, a sparse linear regression (*31*) was applied to predict a neurofeedback signal (i.e., likelihood of the target orientation in V1/V2 during the induction stage) from an activation pattern in each of the 5 areas mentioned above for each trial of the induction stage, and an activation pattern from the V1/V2 itself as a control. Note that the output of the sparse logistic regression decoder (i.e., likelihood of the target orientation in V1/V2; ranging from 0 to 100%) had been computed using a non-linear function (logistic function) (*28*). Thus, before applying the sparse linear regression for each area, the likelihoods in V1/V2 were linearized using an arc hyperbolic tangent function. A predicted value was obtained as the linearly weighted sum of the voxel activities in each area. Prediction accuracy was defined as a coefficient of

determination and evaluated by a leave-one-day-out cross validation procedure. That is, the pair of the V1/V2 likelihoods and the activation patterns for each area measured on one day during the induction stage were treated as the test data while those measured on the remaining days were used for training the sparse linear regression decoder to predict trial-by-trial likelihoods in V1/V2. Five or ten cross-validation sets were generated. A coefficient of determination here indicates the proportion of variability in the likelihoods on a trial-by-trial basis in V1/V2 that is explained by voxel activities in each area. The coefficient of determination for each area was first averaged over the cross-validation sets and then across the subjects and is shown in Fig. S9A. The coefficient for V1/V2 itself was 71% and high, but those for the other 5 areas were les than 5%. The numbers of voxels selected by the sparse algorithm to obtain the predicted values were 371±32 (V1/V2), 203±13 (V3), 259±11 (V4), 629±24 (IPS), and 812±67 (LPFC), respectively.

In the second offline test, we constructed five multinomial sparse logistic regression decoders (*28*) for V3, V4, IPS, LPFC, and the combined region from V3 to LPFC. The entire procedure was the same as described for the fMRI decoder construction stage in the main text. Using the same procedure as for V1/V2, decoding accuracy was evaluated using a leave-one-run-out cross validation procedure for each of the 5 areas. Ten cross-validation sets were generated for each area, and accuracies for the sets were first averaged over the cross-validation sets and then across the subjects. Fig. S9B shows the results of these 5 areas as well as that from V1/V2 as a control. The numbers of voxels selected by the sparse algorithm for decoding were 239±29 (V1/V2), 178±25 (V3), 194±15 (V4), 235±30 (IPS), and 223±13 (LPFC), respectively.

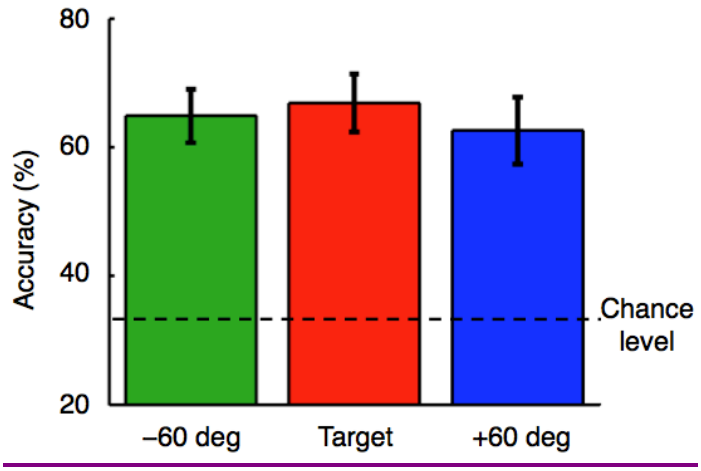

Fig. S1. Performance of a decoder for V1/V2. A leave-one-run-out cross validation procedure (*28*) confirmed that the decoder successfully predicts the Gabor orientations presented to the subjects based on the fMRI datasets measured in the fMRI decoder construction stage. The mean accuracy (±s.e.) for each orientation was significantly higher than the 33% chance level ($t(9)>6.98$, $P<10^{-4}$).

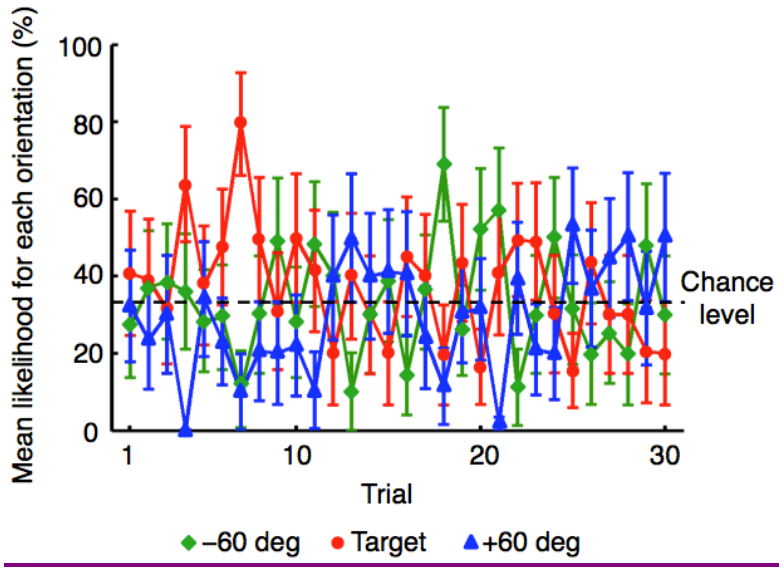

Fig. S2. The mean (±s.e.) likelihood of each of three orientations in the first 30 trials of the first day in the induction stage.

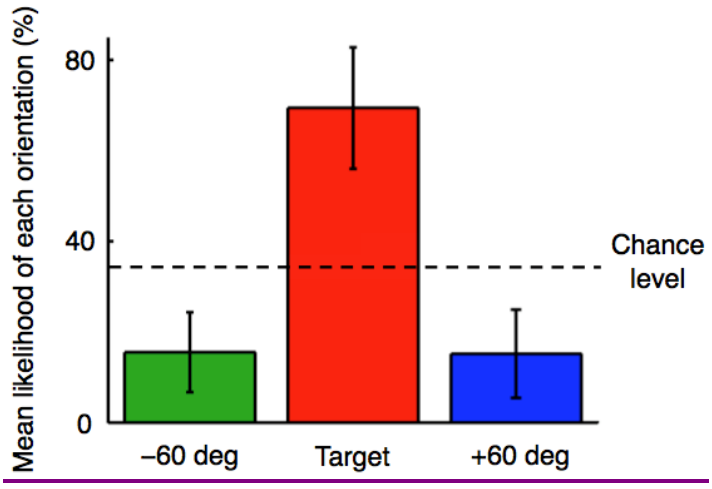

Fig. S3. Results of decoding analysis to the overall mean activation pattern. We also applied the same decoder to the overall mean activation pattern, rather than trial-by-trial activation patterns, in V1/V2 during the induction stage for each subject. Consistent with the results of trial-by-trial decoding, the mean target-orientation likelihood was significantly higher than chance level ($t(9)=2.69$, $P=0.02$).

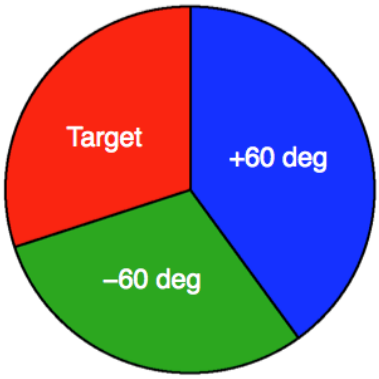

Fig. S4. The percentages of orientations chosen as the target orientation by subjects. The percentages were statistically undistinguishable from what would be expected from chance (Chi-square test, χ^2=0.20, P=0.90).

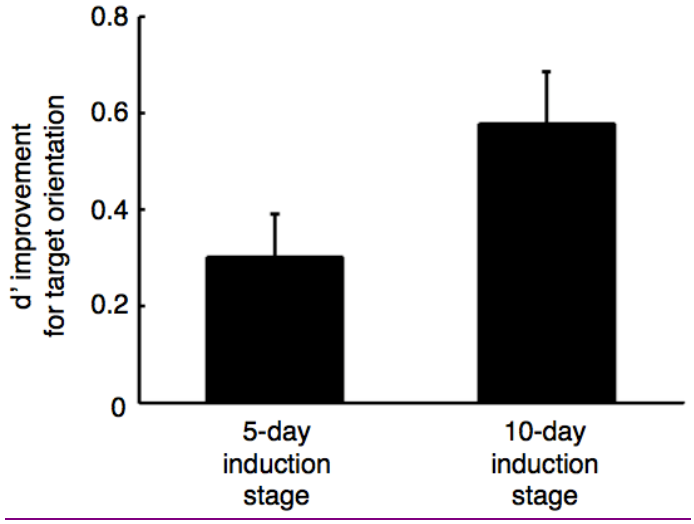

Fig. S5. The mean (±s.e.) d' improvements for the target orientation at 6% S/N ratio. Significant improvements were found for both subjects who participated in 5-day induction stage (left, $t(3)=3.38$, $P=0.04$) and the subjects in 10-day induction stage (right, $t(5)=5.36$, $P<10^{-2}$). No significant difference was found between the two groups (left, $t(8)=1.81$, $P=0.11$).

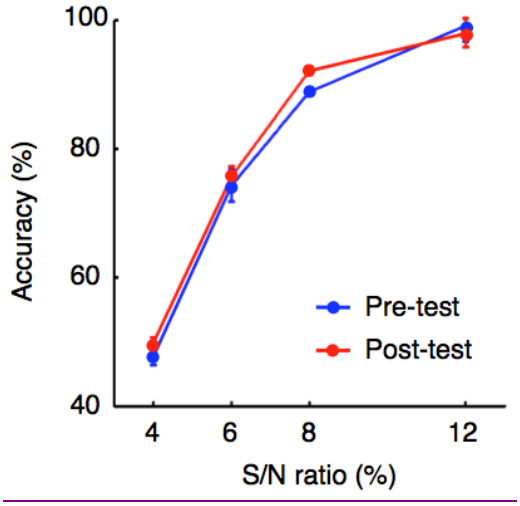

Fig. S6. Results of the pre- and post-tests in the control experiment (n=6) with no induction stage. This plot shows the mean (±s.e.) discrimination accuracies across 10 deg, 70 deg, and 130 deg in the pre- and post-tests as a function of the S/N ratio. Two-way (test stage x S/N ratio) ANOVA with repeated measures indicate no significant effect of test stage ($F(1, 5)=3.25$, $P=0.13$) or interaction of test stage and S/N ratio ($F(3, 15)=1.71$, $P=0.21$).

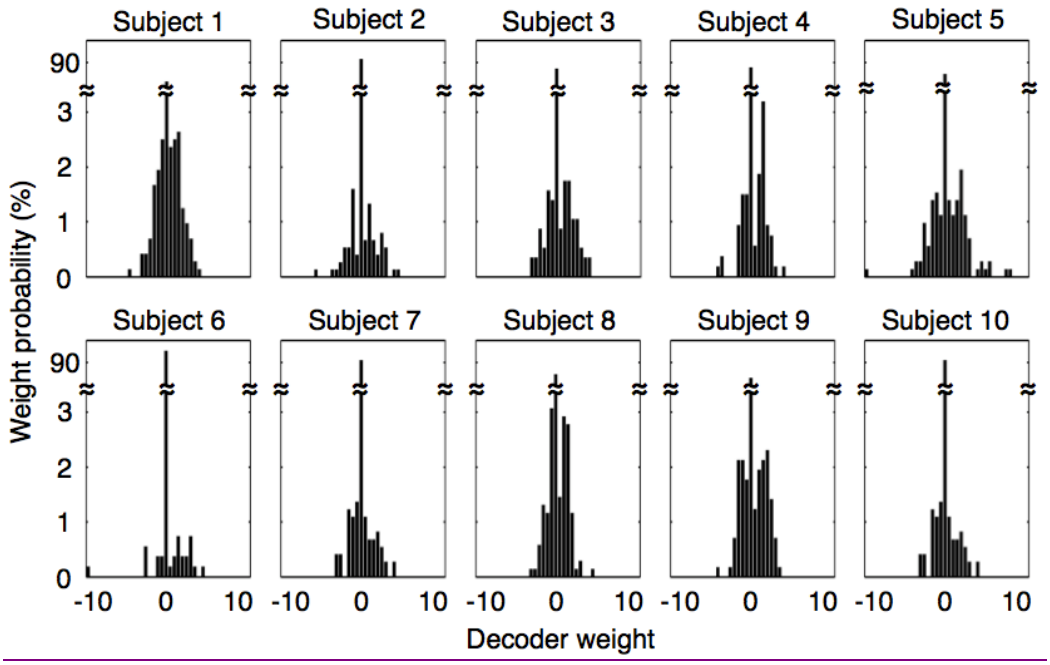

Fig. S7. Histograms of decoder weights for the target orientation in V1/V2 voxels for each subject. Multinomial sparse logistic regression automatically selected relevant voxels, which were informative (decoder weights not being zero) to decode the Gabor orientations, based on fMRI datasets in the fMRI decoder construction stage. The histograms show that about 13% of voxels were assigned to non-zero decoder weights, which were distributed to both positive and negative values. Thus, the high likelihoods computed for the target orientations cannot be obtained simply by increasing global activation in V1/V2.

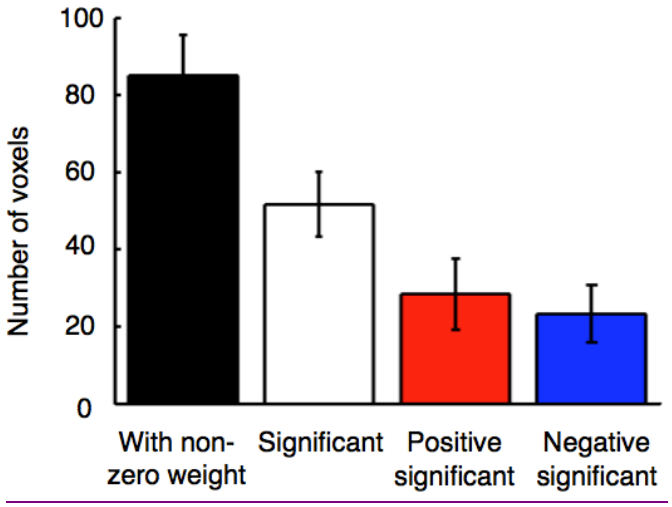

Fig. S8. The mean (±s.e.) number of voxels that showed significant activation during the induction stage. On average, 85 voxels were assigned to non-zero weight (black bar), and 52 of those voxels had significant activation (white bar; t-test, $P<0.05$, corrected by the number of voxels). Approximately, half of the significant voxels were activated positively (red bar) while the other half were activated negatively (blue bar).

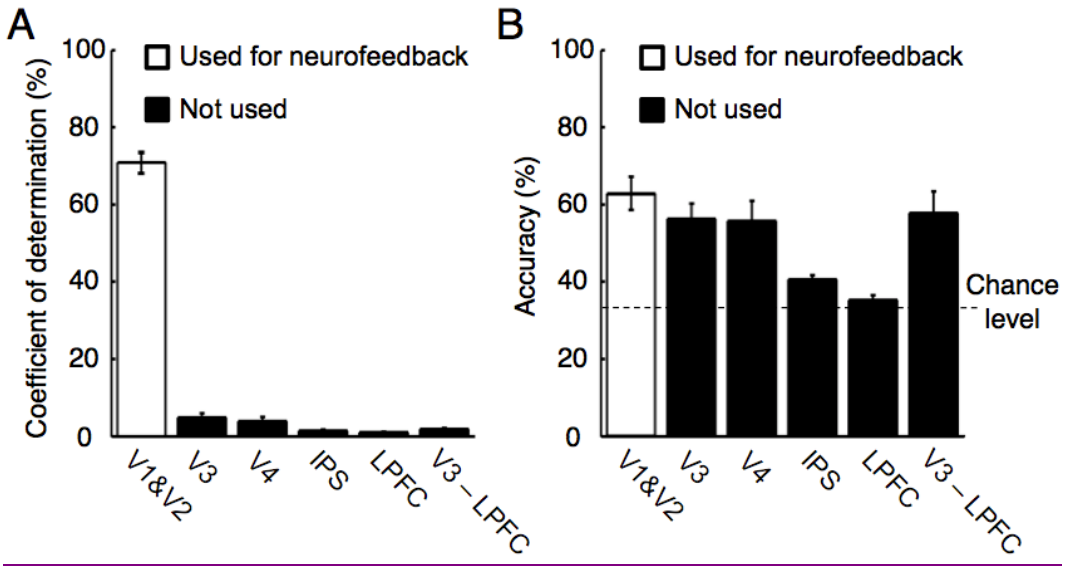

Fig. S9. Results of two offline tests. (A) The mean (±s.e.) coefficient of determination (goodness of fits between the likelihood in V1/V2 and its predicted value for each area, multiplied by 100) for the sparse-linear-regression prediction by each of activation patterns in V3, V4, intraparietal sulcus (IPS), lateral prefrontal cortex (LPFC), and V3-LPFC combined region during the induction stage, and from the V1/V2 itself as a control. The mean coefficients of determination were less than 5% in the 5 areas other than V1/V2. (B) Performance of a multinomial sparse regression decoder for each of the 6 areas including V1/V2 in the fMRI decoder construction stage. Two-way (orientation x area) ANOVA with repeated measures showed significant effect of area ($F(5, 45)=15.01$, $P<10^{-4}$) but no significant effect of orientation ($F(2, 18) = 1.70$, $P = 0.21$) and interaction between area and orientation ($F(10, 90)=1.19$, $P=0.31$) on decoding accuracies. Thus, the accuracies for the 3 orientations were averaged. Results indicated that the decoder successfully predicted the Gabor orientations presented to the subjects based on the fMRI datasets measured in the fMRI decoder construction stage for the areas V1/V2 ($t(9)=6.98$, $P=10^{-4}$), V3 ($t(9)=6.25$, $P < 10^{-3}$), V4 ($t(9)=4.42$, $P<10^{-2}$), IPS ($t(9)=7.50$, $P<10^{-4}$), LPFC ($t(9)=2.69$, $P=0.02$), and V3-LPFC combined region ($t(9)=4.81$, $P<10^{-3}$).

References and Notes

1. A. Schoups, R. Vogels, N. Qian, G. Orban, Practising orientation identification improves orientation coding in V1 neurons. *Nature* **412**, 549 (2001). doi:10.1038/35087601 Medline

2. Y. Yotsumoto, T. Watanabe, Y. Sasaki, Different dynamics of performance and brain activation in the time course of perceptual learning. *Neuron* **57**, 827 (2008). doi:10.1016/j.neuron.2008.02.034 Medline

3. T. Hua *et al*., Perceptual learning improves contrast sensitivity of V1 neurons in cats. *Curr. Biol.* **20**, 887 (2010). doi:10.1016/j.cub.2010.03.066 Medline

4. N. Censor, Y. Bonneh, A. Arieli, D. Sagi, Early-vision brain responses which predict human visual segmentation and learning. *J. Vision* **9**, 1 (2009). doi:10.1167/9.4.12 Medline

5. A. Karni, D. Sagi, The time course of learning a visual skill. *Nature* **365**, 250 (1993). doi:10.1038/365250a0 Medline

6. C. T. Law, J. I. Gold, Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* **11**, 505 (2008). doi:10.1038/nn2070 Medline

7. T. Yang, J. H. Maunsell, The effect of perceptual learning on neuronal responses in monkey visual area V4. *J. Neurosci.* **24**, 1617 (2004). doi:10.1523/JNEUROSCI.4442-03.2004 Medline

8. C. M. Lewis, A. Baldassarre, G. Committeri, G. L. Romani, M. Corbetta, Learning sculpts the spontaneous activity of the resting human brain. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 17558 (2009). doi:10.1073/pnas.0902455106 Medline

9. O. Yamashita, M. A. Sato, T. Yoshioka, F. Tong, Y. Kamitani, Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *Neuroimage* **42**, 1414 (2008). doi:10.1016/j.neuroimage.2008.05.050 Medline

10. S. Bray, S. Shimojo, J. P. O'Doherty, Direct instrumental conditioning of neural activity using functional magnetic resonance imaging-derived reward feedback. *J. Neurosci.* **27**, 7498 (2007). doi:10.1523/JNEUROSCI.2118-07.2007 Medline

11. A. Caria *et al*., Regulation of anterior insular cortex activity using real-time fMRI. *Neuroimage* **35**, 1238 (2007). doi:10.1016/j.neuroimage.2007.01.018 Medline

12. R. C. deCharms *et al*., Learned regulation of spatially localized brain activation using real-time fMRI. *Neuroimage* **21**, 436 (2004). doi:10.1016/j.neuroimage.2003.08.041 Medline

13. N. Weiskopf *et al*., Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): Methodology and exemplary data. *Neuroimage* **19**, 577 (2003). doi:10.1016/S1053-8119(03)00145-9 Medline

14. A. Toda, H. Imamizu, M. Kawato, M. A. Sato, Reconstruction of two-dimensional movement trajectories from selected magnetoencephalography cortical currents by combined sparse Bayesian methods. *Neuroimage* **54**, 892 (2011). doi:10.1016/j.neuroimage.2010.09.057 Medline

15. K. R. Huxlin *et al*., Perceptual relearning of complex visual motion after V1 damage in humans. *J. Neurosci.* **29**, 3981 (2009). doi:10.1523/JNEUROSCI.4882-08.2009 Medline

16. E. Corthout, B. Uttl, V. Walsh, M. Hallet, A. Cowey, Plasticity revealed by transcranial magnetic stimulation of early visual cortex. *Neuroreport* **11**, 1565 (2000). doi:10.1097/00001756-200005150-00039 Medline

17. F. Giovannelli *et al*., Involvement of the parietal cortex in perceptual learning (Eureka effect): An interference approach using rTMS. *Neuropsychologia* **48**, 1807 (2010). doi:10.1016/j.neuropsychologia.2010.02.031 Medline

18. H. R. Dinse, P. Ragert, B. Pleger, P. Schwenkreis, M. Tegenthoff, Pharmacological modulation of perceptual learning and associated cortical reorganization. *Science* **301**, 91 (2003). doi:10.1126/science.1085423 Medline

19. Y. Miyashita, Cognitive memory: Cellular and network machineries and their top-down control. *Science* **306**, 435 (2004). doi:10.1126/science.1101864 Medline

20. M. Kawato, From 'understanding the brain by creating the brain' towards manipulative neuroscience. *Philos. Trans. R. Soc. London Ser. B* **363**, 2201 (2008). doi:10.1098/rstb.2008.2272 Medline

21. A. R. Seitz, D. Kim, T. Watanabe, Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron* **61**, 700 (2009). doi:10.1016/j.neuron.2009.01.016 Medline

22. S. A. Engel, G. H. Glover, B. A. Wandell, Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb. Cortex* **7**, 181 (1997). doi:10.1093/cercor/7.2.181 Medline

23. D. Fize *et al*., The retinotopic organization of primate dorsal V4 and surrounding areas: A functional magnetic resonance imaging study in awake monkeys. *J. Neurosci.* **23**, 7395 (2003). Medline

24. Y. Yotsumoto *et al*., Location-specific cortical activation changes during sleep after training for perceptual learning. *Curr. Biol.* **19**, 1278 (2009). doi:10.1016/j.cub.2009.06.011 Medline

25. Y. Kamitani, F. Tong, Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* **8**, 679 (2005). doi:10.1038/nn1444 Medline

26. S. Nishida, Y. Sasaki, I. Murakami, T. Watanabe, R. B. Tootell, Neuroimaging of direction-selective mechanisms for second-order motion. *J. Neurophysiol.* **90**, 3242 (2003). doi:10.1152/jn.00693.2003 Medline

27. Y. Miyawaki *et al*., Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron* **60**, 915 (2008). doi:10.1016/j.neuron.2008.11.004 Medline

28. O. Yamashita, M. A. Sato, T. Yoshioka, F. Tong, Y. Kamitani, Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *Neuroimage* **42**, 1414 (2008). doi:10.1016/j.neuroimage.2008.05.050 Medline

29. D. H. Brainard, The psychophysics toolbox. *Spat. Vision* **10**, 433 (1997). doi:10.1163/156856897X00357 Medline

30. B. Fischl *et al*., Automatically parcellating the human cerebral cortex. *Cereb. Cortex* **14**, 11 (2004). doi:10.1093/cercor/bhg087 Medline

31. A. Toda, H. Imamizu, M. Kawato, M. A. Sato, Reconstruction of two-dimensional movement trajectories from selected magnetoencephalography cortical currents by combined sparse Bayesian methods. *Neuroimage* **54**, 892 (2011). doi:10.1016/j.neuroimage.2010.09.057 Medline