

Figure S1. Grating stimuli used in the color classifier construction stage, related to Figure 1.

Stimuli used in the color classifier construction stage were red-black, green-black and gray-black gratings whose orientation was either vertical or horizontal.

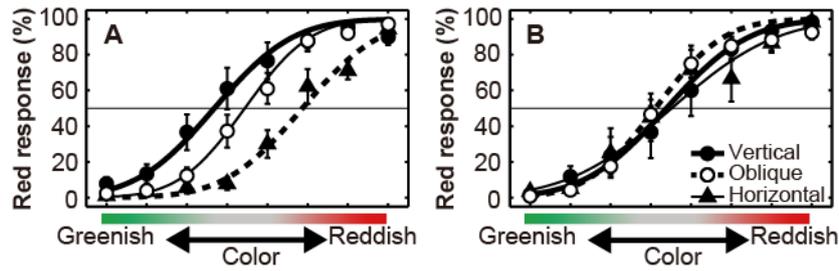


Figure S2. Long-lasting effect of the A-DecNef training on chromatic perception, related to Figure 3.

(**A, B**) Mean (\pm SEM) chromatic psychometric functions measured 3-5 months after induction for the A-DecNef group who were available for the follow-up experiment (**A**, N=9) and for the control group (**B**, N=6, the same data as **Figure 3C**). See **Table S2** for detailed results of three-way mixed-design ANOVA (group as a between-participants factor, and orientation and color as within-participants factors) on chromatic psychometric functions, and one-way repeated measures ANOVA (orientation as a within-participants factor) on the PSE for each orientation.

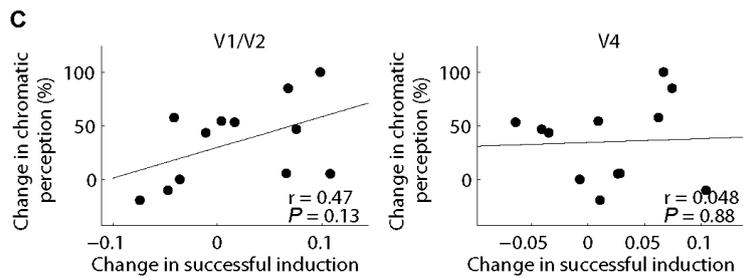
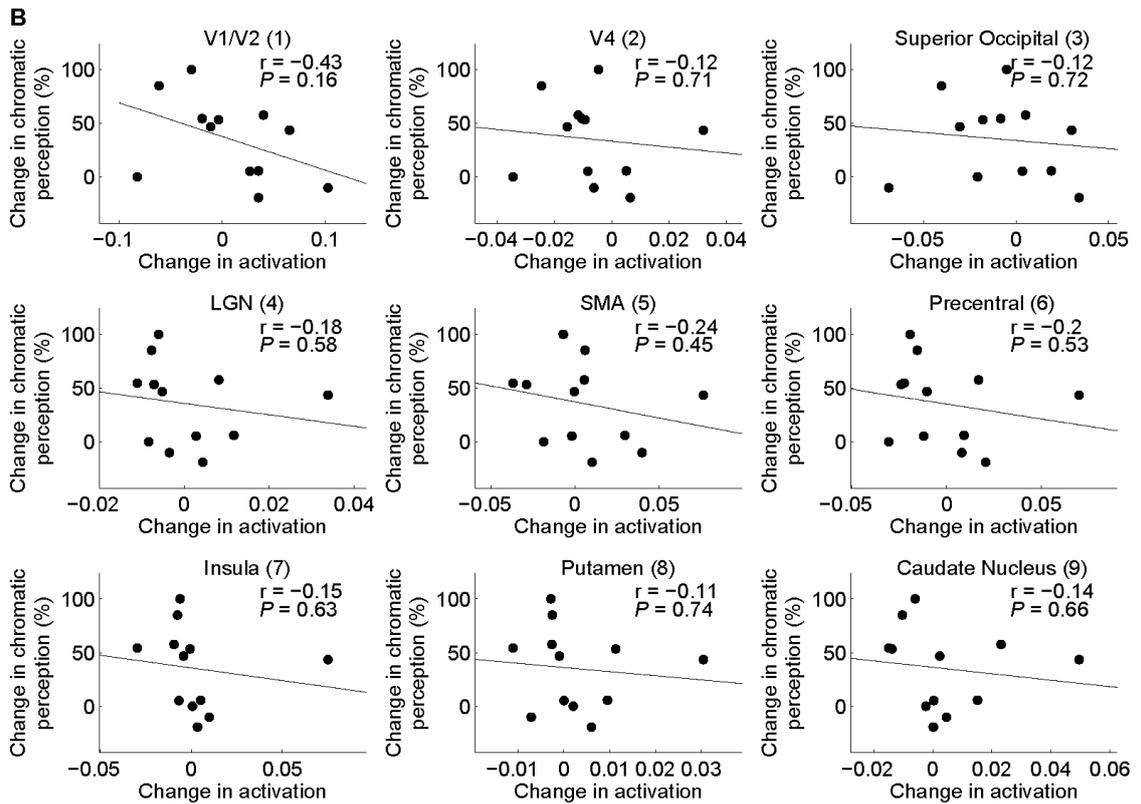
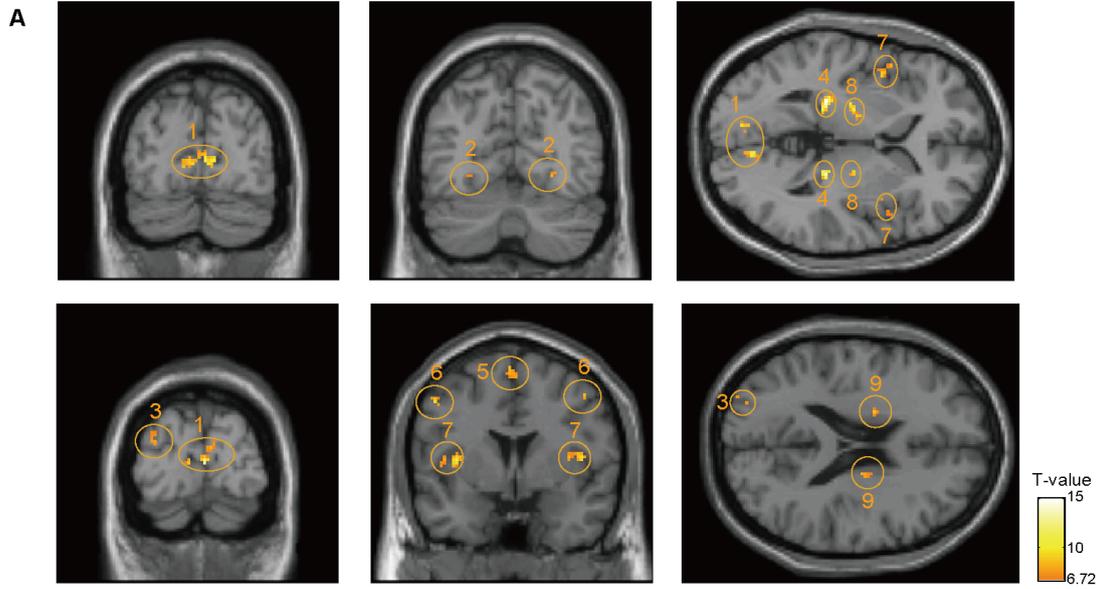


Figure S3. A-DecNef related brain activation, related to Figure 4.

(A) Significant activations ($P < 0.005$) during A-DecNef training (see **Supplemental Experimental Procedures** for details) were found in the nine regions: (1) V1/V2, (2) V4, (3) the superior occipital area, (4) the lateral geniculate nucleus (LGN), (5) the supplementary motor area (SMA), (6) the precentral area, (7) the insula, (8) the putamen, and (9) the caudate nucleus. The names of the areas are according to an anatomy toolbox [S1]. Note that the V4 activation here extends into the fusiform area.

(B) Scatter plots between the chromatic perceptual change and the neural changes, which were estimated using the activation amplitudes in each of the nine regions during A-DecNef training (see **Supplemental Experimental Procedures** for details). The number next to the area name corresponds to the region number shown in (A). The change in the chromatic perception was defined by the sum of the percentage of red response for the vertical grating and that of green response for the horizontal grating in the post-test. The correlation coefficient for V1/V2 was the strongest among the nine regions (see below for more details). One possible reason for the negative correlation for V1/V2 could be that the current associative learning is at least partially due to an overall increase in inhibition. If our model of changes in mutual inhibition is correct (see **Figure S4**), the amount of inhibition from RV (neural population for red & vertical) to GV (neural population for green & vertical) might be larger than the amount of inhibition from GV to RV, which could cause the overall increase in inhibitory signals. Future research is necessary to test this possibility.

(C) Scatter plots between the chromatic perceptual change and the neural changes, which were estimated using the number of successful inductions in V1/V2 and V4

during A-DecNef training (see **Supplemental Experimental Procedures** for details). These two areas were chosen based on their relatively high color classification accuracy (**Figure 4A**). The correlation coefficient for V1/V2 was stronger than V4 (see below for more details). See (**B**) for the change in the chromatic perception.

The GLM analysis showed A-DecNef-related activations in the nine regions. However, based on the correlation analyses with chromatic perceptual changes (**Figure S3B** and **S3C**) as well as the results of the searchlight analysis (**Figure 4**), we suggest that not all of A-DecNef related activations are involved in the change of the red likelihood (**Figure 2**). The results as in the **Figure S3B** shows that only V1/V2 showed the moderate strength of correlation ($r = -0.43$), although not statistically significant, whereas other areas showed little correlation (correlation coefficients ranging $r = -0.24$ to -0.11) with the chromatic perceptual changes. This suggests that V1/V2 is most related to the change in color perception among these nine regions. While the searchlight analyses (**Figure 4**) suggested that the red likelihood may be represented in V4 as well as in V1/V2, depending on the alpha level, the results as in **Figure S3C** indicate that V1/V2 has stronger correlation coefficient, again supporting the idea that V1/V2 is likely to be the main locus for the association between orientation and color.

We speculate that A-DecNef-related activations may correspond to visual processing and skill learning in addition to the red likelihood information, which is likely to be represented in V1/V2. We will describe our interpretations of activations other than V1/V2 below. First, we speculate that the significant activations in the LGN, V4 and the superior occipital area correspond to visual processing since a visual stimulus (achromatic grating) was presented during A-DecNef. Second, the significant

activations in the remaining regions (the SMA, the precentral gyrus, the insula, the putamen and the caudate nucleus) have been implicated in motor sequence or skill learning (for instance, [S2-5]). We speculate that A-DecNef involves skill learning because participants learn to modulate their brain activations. Interestingly, in a brain-computer interface (BCI) study [S6], the SMA has been shown to be involved. BCI is a neurorehabilitation technology [S7] that requires participants to learn to move external devices without using normal motor output (for instance, to move a cursor which receives the motor command from the motor cortex without using a hand). Thus, participants have to learn to modulate activation of their motor cortex. A-DecNef also required a participant to learn to change activation patterns of the early visual cortex (although it was not clear to the participant). Regarding the activations of the putamen and caudate nucleus, alternative interpretation may be possible, because they have been shown to be involved in instrumental conditioning [S8]. Since A-DecNef can be considered to be an instrumental conditioning, the activation of putamen and caudate nucleus may correspond to an instrumental conditioning component. Together, the present A-DecNef recruits areas including those involved in visual and motor functions, whose roles in A-DecNef have yet to be revealed.

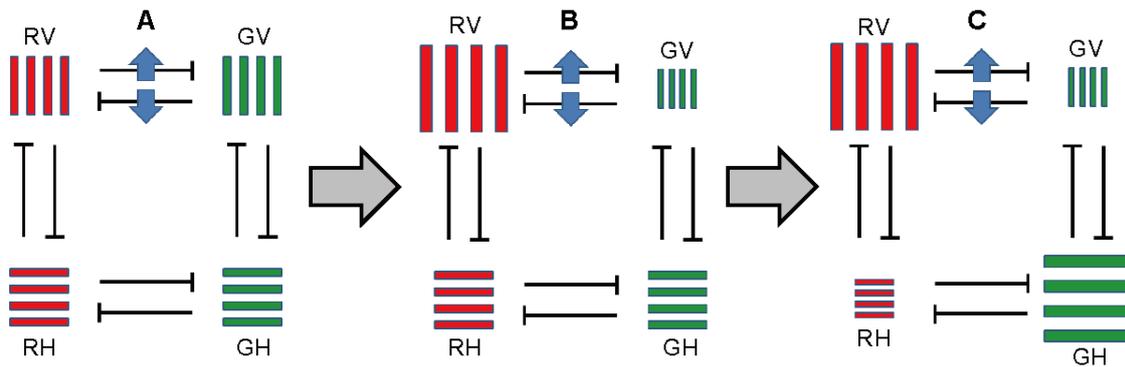


Figure S4. Possible neural mechanisms underlying the change in chromatic perception by A-DecNef, related to Figure 3.

Here we assume four populations for the sake of simplicity. Each population has sensitivity toward both color (red vs. green) and orientation (vertical vs. horizontal). They could jointly be sensitive to both color and orientation, as binding features [S9, 10], or they could contain two sub-populations each of which has sensitivity towards either color or orientation alone but they are associated. The assumed four populations are red-vertical (RV), green-vertical (GV), red-horizontal (RH) and green-horizontal (GH) gratings. We propose that the populations mutually inhibit between opposing values of color (red vs. green) or orientation (vertical vs. horizontal).

(A) In the model, first, A-DecNef both increases inhibition from RV to GV and decreases inhibition from GV to RV. (B) Second, the imbalance in mutual inhibition results in larger response of RV than GV, which leads to reddish perception on the vertical grating (C) Third, inhibition from increased response of RV to RH leads to decreased response of RH, whereas inhibition from decreased response of GV to GH leads to increased response of GH. Larger response of GH than RH leads to greenish perception on the horizontal grating.

Table S1, related to Figure 1. At the end of each A-DecNef session, participants were queried regarding any techniques, strategies, or particular thoughts they employed to increase the size of the feedback disk. Responses were as follows (Japanese responses were translated into English by the authors).

Participant	Day1	Day2	Day3
1	“I concentrated on the fixation point and attended to the vertical grating.”	“I imagined a situation where I behaved violently.”	“I sometimes tried to focus on, or at other times defocused from, the stimulus.”
2	“I tried to concentrate on the task sometimes, or relaxed at other times. I also rotated the attended area slowly (several seconds per rotation) while maintaining fixation at the center.”	“I imagined that the black area around the fixation point was contracting. I also moved the attended location in a circular trajectory.”	“I moved the attended location in a circular trajectory.”
3	“I tried to remember scenes shown on TV or video games. I also imagined watching music videos.”	“I tried to remember scenes in video games. Then I imagined music of promotional films.”	“I tried to remember scenes in video games.”
4	“I attended to the anterior part of the brain. Then I tried to control activity in the occipital area. I also focused attention on the fixation point.”	“I focused attention on the grating.”	“I focused attention on the grating.”
5	“I imagined that I was singing or running. I also made calculations	“I imagined visual and auditory images”	“I imagined scenes in promotional films.”

	on some things or remembered some sentences.”		
6	“I tried to remember what I did yesterday. I also tried to recall memories of exciting incidences during my last summer vacation.”	“I tried to remember what I was impressed with.”	“I tried to remember memories of exciting incidences during my last summer vacation.”
7	“I did mental multiplication. I also imagined several things including a zebra and building. I also imagined a uniform gray circle.”	“I imagined a uniform gray circle.”	“I imagined a zebra. I also imagined a disk.”
8	“I imagined myself to be a hero. I also imagined actions including lighting a cigarette or getting in a car.”	“I imagined actions including lighting a cigarette or getting in a car. I also mentally counted, or remembered the lyrics of songs. ”	“I imagined actions including lighting a cigarette or getting in a car. I also imagined the movements of a bus driver.”
9	“I first calculated, and then remembered what I did yesterday, and the day before yesterday. I also planned what I will do today.”	“I added up figures in my head.”	“I added up figures in my head.”
10	“I just focused my attention on the fixation point.”	“I imagined translational movement of an afterimage. I also thought about something.”	“I mentally drew a circle within the black region around the fixation point. I also imagined translational movement of an afterimage or mentally let out a

			voice.”
11	“I imagined a gymnastic match in which I performed well. I also imagined simple actions including running.”	“I imagined a gymnastic match in which I performed well. I also imagined simple actions including running.”	“I imagined simple actions including running or bicycling.”
12	“I mentally moved a dot along the annulus aperture of the grating.”	“I mentally moved a dot along the annulus aperture of the grating.”	“I mentally moved a dot along the annulus aperture of the grating.”

Table S2. ANOVAs for long-term effect of A-DecNef, related to Figure 3. (A)

Results of three-way ANOVA on the red response%. (B) Results of one-way ANOVA on the PSE for each orientation. (C) Results of t-test on the red response% at the neutral gray.

A: Three-way mixed-design ANOVA (group as a between-participants factor, and orientation and color as within-participants factors) on the red response%

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
orientation	14821.20	1.2	12505.30	8.59	.008
orientation x Group	13234.54	1.2	11166.56	7.67	.011
Error (orientation)	22436.57	15.4	1456.21		
Color	415965.19	4.4	94510.58	146.69	<.001
color x Group	1836.30	4.4	417.22	.648	.645
Error (color)	36863.43	57.2	644.28		
orientation x color	9908.43	8.7	1136.64	3.86	<.001
orientation x color x Group	10330.65	8.7	1185.07	4.03	<.001
Error (orientation x color)	33367.13	113.3	294.44		

Simple interaction between orientation and color for each group of participants

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
orientation x color (A-DecNef)	20825.93	14	1487.57	8.11	<0.001
orientation x color (control)	2981.94	14	213.00	1.16	0.308
Error (orientation x color)	33367.13	182	183.34		

B: One-way repeated measures ANOVA (orientation as a within-participants factor) on the PSE for each orientation

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Orientation	19.34	1.2	15.53	20.94	0.001
Error (orientation)	7.39	10.0	0.74		

Multiple comparisons for the PSE

Pair	df	t	Sig.
Vertical vs. Horizontal	8	-4.74	0.002
Vertical vs. Oblique	8	-3.11	0.015
Horizontal vs Oblique	8	5.56	<0.001

C: T-test on the red response% at the neutral gray

Source	df	t	Sig.
Vertical vs 50%	8	4.22	0.003
Horizontal vs 50%	8	-7.84	<0.001

Supplemental Experimental Procedures

Participants

Data from 18 participants (18 to 27 years old) are included in the present experiments. Nine of the 12 participants took part in the experiment that tested the long-lasting effect. Another six participants took part in the control experiment, which consisted only of the post-test stage, without A-DecNef training. All participants had normal or corrected-to-normal color vision and gave their written informed consent to participate. All experimental procedures were approved by the institutional review board at Advanced Telecommunications Research Institute International (ATR).

The experimental design for the main experiment

The main experiment consisted of four stages: (1) retinotopic mapping (one day), (2) color classifier construction (one day), (3) A-DecNef training (three days), and (4) post-test stages. The color classifier construction stage and the first day of the A-DecNef training stage were separated by at least ten days. The post-test was conducted immediately after the conclusion of the A-DecNef training stage.

MRI parameters

Participants were scanned in a 3T MRI scanner (Verio, Siemens) with a head coil at the ATR Brain Activation Imaging Center. fMRI signals were acquired using a gradient EPI sequence. In all fMRI experiments, 33 contiguous slices ($TR = 2$ sec, voxel size = $3 \times 3 \times 3.5$ mm³, 0 mm slice gap) oriented parallel to the AC-PC plane were acquired, covering the entire brain. T1-weighted MR images (MP-RAGE; 256 slices, voxel size = $1 \times 1 \times 1$ mm³, 0 mm slice gap) were also acquired.

Apparatus and stimuli

In the MRI scanner, visual stimuli were presented using an LCD projector (DLA-G150CL, Victor) on a translucent screen. Measurement of chromatic psychometric functions in the post-test stage was conducted outside the scanner using a CRT display (P275, IBM). Both the projector and CRT spanned 20×15 deg of visual angle (800×600 resolution) and had a 60 Hz refresh rate. All visual stimuli were generated using Psychtoolbox 3 [S11] running on Matlab.

Retinotopic mapping

Before the color classifier construction stage, we measured retinotopic maps [S12] using conventional rotating wedge and expanding ring stimuli presented within a circular region of 7.5 deg in radius, allowing us to define V1, V2, and V4 for each participant. In addition, participants were presented with a localizer stimulus to define the subregions (1.5-7.5 deg in eccentricity) of V1, V2, and V4. The localizer stimulus was a flickering colored checkerboard, which was presented alternatively between one visual field (an annulus region; 1.5-7.5 deg in radius) for 8 sec and another (a central circle region; 0-1.5 deg in radius) for 8 sec. The 16 sec block was repeated 14 times per run for at least 3 runs. To ensure the maintenance of gaze at the fixation point presented at the center of the screen, during both the retinotopic mapping and localizer scans, participants were asked to press a button with their right hand if they detected changes in the color of the fixation point.

Color classifier construction stage

In the color classifier construction stage, we measured BOLD-signal multi-voxel patterns evoked by the presentation of red-black, green-black, and gray-black gratings oriented both vertically and horizontally (**Figure S1**). The grating stimulus, presented within an annular aperture (1.5-7.5 deg in radius), was defined by a square-wave modulation between either red, green, or gray and black at 1 cycle/deg. The phase of the grating was fixed across all trials, runs, and participants, so that the position of the black segments of the grating was constant. To construct a classifier based on color difference rather than on luminance difference, the red, green and gray colors were adjusted to be iso-luminant to each other. The red luminance was fixed at the maximum output of the red gun ([255 0 0], $x=0.643$, $y=0.326$, $Y=22.3$ in

CIE xyY color space). The green luminance ([0 Green 0]) and the gray luminance ([Gray Gray Gray]) were matched to the red luminance for each participant using flicker photometry [S13]. During flicker photometry, red/green or red/gray flicker patterns were presented at 30 Hz, and participants were instructed to minimize flicker by adjusting green/gray luminance with a keyboard. Measurements were repeated ten times for each color, and the resulting green and gray colors, perceptually iso-luminant to red, were used for the color classifier construction stage.

In each trial of the color classifier construction stage, the grating flickered at 0.5 Hz for 6 sec (3 repetitions of 1.5 sec on and 0.5 sec off) and was followed by a black screen for 6 sec. The color (red, green, or gray) and orientation (vertical or horizontal) of the grating were selected randomly from the six possible combinations. For the purpose of maintaining fixation, participants performed a fixation task in which they were asked to press a button with their right hand if they detected changes in the luminance (from white to gray) of the fixation point at the center of the screen. The difficulty of the fixation task was manipulated with a QUEST procedure [S14] so that the hit ratio asymptotically approached 50%.

Within a run, each combination of color and orientation was presented 4 times. Ten sec and 6 sec fixation periods were placed before and after the stimulus presentations, respectively (1 run = 304 sec). Twelve runs were repeated for each participant. A brief break period was provided after each run upon a participant's request.

Measured BOLD signals were preprocessed using the mrVista software developed at Stanford University (<http://vistalab.stanford.edu/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 participant. A gray matter mask was used to extract BOLD signals only from gray matter voxels for further analyses. A reference region within V1/V2 that corresponded to the size of the grating stimulus (1.5-7.5 deg in radius) was defined by intersecting V1/V2 and the area activated by the localizer. Once we identified the reference region, the time-course of BOLD signals in the color classifier construction stage was extracted from each voxel in that region. After removing the linear trend, the time-course was z-score normalized for each voxel for each run to minimize baseline differences across the runs. The data samples for computing the classifier were created by averaging the BOLD-signal amplitudes of each voxel across the 3 TR corresponding to the 6 sec grating periods. Six sec (3 TR) of hemodynamic delay was taken into account.

We used sparse logistic regression (SLR) [S15] to construct a red-black vs. green-black (red vs. green) classifier for live use during the A-DecNef training stage. SLR automatically selected the voxels relevant for the discrimination of grating color from the voxels in the reference region within V1/V2. We trained the classifier to classify BOLD-signal multi-voxel patterns into one of the two colors using 192 data samples obtained from 192 trials during the 12 runs. The classifier did not contain orientation information because BOLD-signal multi-voxel patterns from both vertical and horizontal gratings were pooled for the analysis. The mean (\pm SEM) number of voxels in V1/V2 selected by SLR was 194 ± 28 .

After we built the classifier, we also calculated the target color (red) likelihood based on the BOLD-signal multi-voxel patterns evoked by the gray-black grating. This likelihood is shown as on Day 0 in **Figure 2**.

In a separate analysis, we tested whether two different colors can be differentiated based on different BOLD-signal multi-voxel patterns in V1/V2. A leave-one-run-out cross-validation procedure was conducted for each participant. In each cross-validation cycle, the classifier was trained using 176 data samples obtained from 11 runs and tested using 16 data samples obtained from the remaining run. This cycle was repeated 12 times. The mean classification accuracy for the test data across the 12 cycles was regarded as decoding accuracy.

A-DecNef training stage

During the A-DecNef training stage, participants were trained to create an internal association between the presentation of an achromatic vertical grating and the neural activity corresponding to a specific target color (red).

A-DecNef training stage always spanned 3 consecutive days (1 session per day). Each day consisted of a maximum of 12 runs. The mean (\pm SEM) number of runs on each day was 11.9 ± 0.1 across days and participants. A brief break period after each run was provided upon a participant's request to minimize fatigue.

Each run was 330 sec long and consisted of 15 trials (1 trial = 20 sec) preceded by a 30 sec fixation period. Each trial consisted of an induction period (6 sec), a fixation period (7 sec), a feedback period (1 sec), and an inter-trial interval (6 sec), in that order (**Figure 1**). During the induction period, an achromatic (gray-black) vertical grating, identical to the one used in the color classifier construction stage (**Figure S1**), was presented with a fixation point at the center of the screen. The grating flickered at 0.5 Hz for 6 sec (3 repetitions of 1.5 sec on and 0.5 sec off). During the fixation period and inter-trial interval, only a fixation point was presented on a black background. Participants were instructed to somehow regulate their brain activity during the presentation of the achromatic grating to make the size of the solid gray disk presented during the subsequent feedback period as large as possible. Participants were also instructed to maintain fixation on the fixation point for the whole experiment period. The experimenters provided no further instructions or strategies.

The size of the gray disk presented during the feedback period represented how much the BOLD-signal multi-voxel patterns in V1/V2 during the induction period corresponded to the patterns evoked by the presentation of the red-black gratings in the color classifier construction stage. The size of the gray disk was determined by the following procedures. First, newly obtained functional image data underwent 3D head motion correction using Turbo-Brain Voyager (Brain Innovation, The Netherlands). Second, time-course of BOLD signals was extracted from each voxel in the reference region within V1/V2. Third, a linear trend was removed from the time-course. The time-course was then z-score normalized for each voxel using a 20 sec time-course of BOLD-signal amplitudes collected starting from 10 sec after the start of each run. Fourth, the BOLD signals measured during the first 6 sec of the fixation period were averaged for each voxel, so that the resultant BOLD-signal multi-voxel patterns would reflect neural activities that were induced during the 6 sec of the induction period (hemodynamic delay was assumed to be 6 sec). Finally, the red likelihood was calculated by multiplying the BOLD-signal multi-voxel patterns with the weights determined in the color classifier construction stage, and by passing this weighted sum through a logistic function (likelihood estimates could range from 0 to 100%). The size of the gray disk was proportional to the red likelihood and the disk was always enclosed by a circle (1.5 deg in radius) that represented the disk's maximum size.

At the end of each run, the amount of the monetary bonus reward for that run, as well as the accumulated bonus amount, was shown on the screen. The amount of the reward for each run was proportional to the size of the feedback disk averaged across trials. Participants were informed that they would receive bonus rewards proportional to the total size of the disk, but were not informed as to how the size of the disk was determined or what the size represented. Participants were paid a maximum bonus of up to 3,000 JPY per day, in addition to a fixed amount for participation in the experiment.

Immediately after the conclusion of each day of the A-DecNef training stage, participants were asked about how they had manipulated their brain activity (**Table S1**).

Post-test stage

A chromatic psychometric function was measured outside the scanner. In each trial of the psychometric function measurement, a grating of one of four orientations (vertical, horizontal and 2 oblique orientations, **Figure 3A**) in one of eight possible colors (32 configurations in total) was presented. The grating was a square-wave modulation between a color and black at 1 cycle/deg. The color in the inner gratings varied in 8 steps between a reddish tint ($x=0.323$, $y=0.310$) and greenish tint ($x=0.313$, $y=0.323$), passing through a neutral gray, keeping the luminance constant ($Y=17.9$). The outer grating was a square-wave modulation between gray and black at 1 cycle/deg. The phase of the grating was the same as the phase of the grating used in the color classifier construction stage (**Figure S1**). Participants were asked to report whether the inner grating was tinted red or green. Gratings were presented for 0.5 sec, and each combination of orientation and color was repeated 10 times in a random order. A next trial started immediately after participant's response.

The psychophysical experiment was not conducted prior to the A-DecNef training stage. This is because we wished to avoid the possibility that participants would notice the relation between the color and orientation, and explicitly associate color with orientation during the induction periods in the subsequent A-DecNef training stage.

Searchlight analysis

We tested whether other areas than V1/V2 contributed to creation of the association of orientation-specific color perception. We conducted a searchlight analysis [S16] by moving a sphere ($r=15$ mm) region of interest (ROI) across the whole gray matter voxels (~30000 voxels). We calculated (1) color classification accuracy during the color classifier construction stage (**Figure 4A**), (2) the predictability of the red likelihood of V1/V2 by BOLD-signal patterns in other areas than V1/V2 during the color classifier construction stage (**Figure 4B**), and (3) the predictability of the red likelihood of V1/V2 by BOLD-signal patterns in other areas than V1/V2 during the A-DecNef training stage (**Figure 4C**), using the searchlight method [S16].

First, for color classification accuracy, we constructed a color classifier for each sphere ROI using the BOLD-signal patterns measured in the color classifier construction stage and evaluated the classifier performance (percent correct) using the leave-one-run-out cross-validation procedure. The same leave-one-run-out cross-validation procedure was used as for the color classifier construction stage (see above), except that linear support vector machine rather than SLR was employed for its time efficiency. The searchlight analysis demonstrated that both V1/V2 and ventral areas including V4 enjoyed high classification accuracy of color (**Figure 4A**).

Second, the predictability of the red likelihood of V1/V2 by BOLD-signal patterns in each sphere ROI during the color classifier construction stage was computed as below. The red likelihood ranging from 0 to 100%, which was computed by SLR from V1/V2 BOLD-signal patterns (see Color classifier construction stage section), was predicted by L_1 -regularized least-square regressor ($\alpha=1$, $\lambda=0.024$) from BOLD-signal patterns within each sphere ROI. A hyper parameter λ , which decides the balance between the residual and L_1 norm of the weights, was defined by the geometric mean of the optimal λ across three sphere ROIs (centered at V1/V2, left V4 and right V4) for all of the participants (36 ROIs in total). Prediction accuracy was defined as a coefficient of determination between the likelihoods in V1/V2 and the likelihoods calculated using BOLD-signal patterns in the sphere ROI on a trial-by-trial basis. The accuracy was evaluated by a leave-one-run-out cross-validation procedure. The pairs of the red likelihood on one run were used for test data, while the pairs obtained on the remaining runs were used for training the sparse linear regression model.

Third, the predictability of the red likelihood of V1/V2 by BOLD-signal patterns in each sphere ROI during A-DecNef training stage was computed. The L_1 -regularized least-square regressor model of V1/V2 red likelihood was built using the data during the classifier construction stage as described above, and was applied to the data during the A-DecNef training stage consisting of three days. The accuracy of prediction was again evaluated as a coefficient of determination between the actual red likelihood within V1/V2 and its regressed value from a sphere ROI.

The above analyses using the searchlight method were calculated in the native coordinate of each participant, normalized into a MNI template and were averaged across participants. A permutation test was conducted on the classification accuracy (**Figure 4A**) and on the correlation coefficient (**Figures 4B and 4C**) [S17]. P -values were corrected for multiple comparisons using threshold-free cluster enhancement [S18]. The values significantly above chance ($P<0.05$) were plotted.

GLM analysis

We conducted a conventional GLM analysis to find areas activated during the A-DecNef training stage (**Figure S3**) using the SPM8 package [S19]. We calculated the contrast between the induction period (6 s) and the baseline that was automatically computed in the BOLD signal time course by the SPM. The group analysis was performed on the data of all 12 participants, and the areas showing significant activation ($P<0.005$, corrected for family wise error) were identified based on the Gaussian random field theory.

Correlation analysis

For the purpose of elucidating the brain area contributing to the orientation specific chromatic perception, we calculated the correlation coefficients between the change of chromatic perception due to A-DecNef and two types of changes in neural activities (**Figure S3**).

As the change of chromatic perception, we calculated the sum of the percentage of red response for the vertical grating and that of green response for the horizontal grating in the post-test stage.

As for the measure of neural changes, we have 2 ways. We utilized change in activation amplitude for the first measure, and the number of successful induction (the red likelihood higher than 90%) across all 36 runs in three days of A-DecNef training for the second measure. First, we calculated the slope of the best fitted line to the estimated BOLD-signal amplitude (beta value) during the induction period of 36 runs in the 9 ROIs identified by the GLM analysis (**Figure S3A**), and obtained correlation coefficients with the chromatic perceptual change for each of 9 regions (**Figure S3B**). Second, we calculated the slope of the best fitted line to the number of successful inductions of 36 runs in V1/V2 and V4, which showed relatively high color classification accuracy (**Figure 4A**), and obtained correlation coefficients with the perceptual changes for V1/V2 and V4 (**Figure S3C**).

Supplemental References

- S1. Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., and Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325-1335.
- S2. Hikosaka, O., Sakai, K., Miyauchi, S., Takino, R., Sasaki, Y., and Putz, B. (1996). Activation of human presupplementary motor area in learning of sequential procedures: a functional MRI study. *J Neurophysiol* 76, 617-621.
- S3. Doyon, J., and Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. *Curr Opin Neurobiol* 15, 161-167.
- S4. Mutschler, I., Schulze-Bonhage, A., Glauche, V., Demandt, E., Speck, O., and Ball, T. (2007). A rapid sound-action association effect in human insular cortex. *PLoS One* 2, e259.
- S5. Segawa, J.A., Tourville, J.A., Beal, D.S., and Guenther, F.H. (2015). The neural correlates of speech motor sequence learning. *J Cogn Neurosci* 27, 819-831.
- S6. Kasahara, K., DaSalla, C.S., Honda, M., and Hanakawa, T. (2015). Neuroanatomical correlates of brain-computer interface performance. *Neuroimage* 110, 95-100.
- S7. Shindo, K., Kawashima, K., Ushiba, J., Ota, N., Ito, M., Ota, T., Kimura, A., and Liu, M. (2011). Effects of neurofeedback training with an electroencephalogram-based brain-computer interface for hand paralysis in patients with chronic stroke: a preliminary case series study. *J Rehabil Med* 43, 951-957.
- S8. O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452-454.
- S9. Seymour, K., Clifford, C.W., Logothetis, N.K., and Bartels, A. (2010). Coding and binding of color and form in visual cortex. *Cereb Cortex* 20, 1946-1954.
- S10. Zhang, X., Qiu, J., Zhang, Y., Han, S., and Fang, F. (2014). Misbinding of color and motion in human visual cortex. *Curr Biol* 24, 1354-1360.
- S11. Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat Vis* 10, 433-436.
- S12. Wandell, B.A., and Winawer, J. (2011). Imaging retinotopic maps in the human brain. *Vision Research* 51, 718-737.
- S13. Ives, H.E. (1912). Studies in the photometry of lights of different colours. I. Spectral luminosity curves obtained by the equality of brightness photometer and flicker photometer under similar conditions. *Philos. Mag. Ser. 6* 24, 149-188.
- S14. Watson, A.B., and Pelli, D.G. (1983). QUEST: a Bayesian adaptive psychometric method. *Percept Psychophys* 33, 113-120.
- S15. Yamashita, O., Sato, M.A., Yoshioka, T., Tong, F., and Kamitani, Y. (2008). Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *NeuroImage* 42, 1414-1429.
- S16. Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America* 103, 3863-3868.
- S17. Nichols, T.E., and Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp* 15, 1-25.
- S18. Smith, S.M., and Nichols, T.E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44, 83-98.
- S19. Nichols, P.F.A.K. (2006). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*.