Volitional modulation of higher-order visual cortex alters human perception

Article

Accepted Version

Ekanayake, J., Ridgway, G. R., Winston, J. S., Feredoes, E., Razi, A., Koush, Y., Scharnowski, F., Weiskopf, N. and Rees, G. (2019) Volitional modulation of higher-order visual cortex alters human perception. NeuroImage, 188. pp. 291-301. ISSN 1053-8119 doi: https://doi.org/10.1016/j.neuroimage.2018.11.054 Available at http://centaur.reading.ac.uk/81314/

To link to this article DOI: http://dx.doi.org/10.1016/j.neuroimage.2018.11.054

Publisher: Elsevier

TITLE:

Volitional Modulation of Higher-order Visual Cortex Alters Human Perception

AUTHORS:

Jinendra Ekanayake[1,2], Gerard R Ridgway[3], Joel S Winston[1,2], Eva Feredoes[4], Adeel Razi[1,9], Yury Koush[10], Frank Scharnowski[5,6,7], Nikolaus Weiskopf[1,8], Geraint Rees[1,2]

AFFILIATIONS:

[1]Wellcome Trust Centre for Human Neuroimaging, UCL Institute of Neurology, London, UK

[2]UCL Institute of Cognitive Neuroscience, University College London

[3]Oxford Centre for Functional MRI of the Brain, University of Oxford, Oxford, UK

[4]School of Psychology and Language Sciences, University of Reading, UK

[5]Psychiatric University Hospital, University of Zürich, Lenggstrasse 31, 8032 Zürich, Switzerland

[6]Neuroscience Center Zürich, University of Zürich and Swiss Federal Institute of Technology, Winterthurerstr. 190, 8057 Zürich, Switzerland

[7]Zürich Center for Integrative Human Physiology (ZIHP), University of Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland

[8]Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[9]Department of Electronic Engineering, NED University of Engineering and Technology, Karachi, Pakistan

[10]Department of Radiology and Biomedical Imaging, Yale University, 300 Cedar Street, New Haven, CT 06519, USA

CORRESPONDING AUTHOR:

Jinendra Ekanayake, email: j.ekanayake@ucl.ac.uk

ADDRESS:

Institute of Cognitive Neuroscience,
12 Queen Square, WC1N 3AR

Conflict of Interest: Nil

Abstract

Can we change our perception by controlling our brain activation? Awareness during binocular rivalry is shaped by the alternating perception of different stimuli presented separately to each monocular view. We tested the possibility of causally influencing the likelihood of a stimulus entering awareness. To do this, participants were trained with neurofeedback, using realtime functional magnetic resonance imaging (rt-fMRI), to differentially modulate activation in stimulus-selective visual cortex representing each of the monocular images. Neurofeedback training led to altered bistable perception associated with activity changes in the trained regions. The degree to which training influenced perception predicted changes in grey and white matter volumes of these regions. Short-term intensive neurofeedback training therefore sculpted the dynamics of visual awareness, with associated plasticity in the human brain.

Word count 123 (150 max)

Highlights

- Unconscious biasing of higher-order visual perception with realtime fMRI neurofeedback.
- Participants unknowingly modulated two brain regions to control feedback signal
- Short-term neurofeedback training over 3 days induced functional plasticity
- Neurofeedback may strengthen neural representations and alter prior expectations
- Potential avenue for behavioural shaping and therapeutic reduction of aberrant perception

Introduction

The ability to causally modify how we perceive the world has potential implications in health and disease. Altering perceptual biases, which may be conscious or unconscious, could modify pathological perception such as hallucinations, or provide a means of selective cognitive enhancement(Miranda et al., 2015). Such attempts to deliberately manipulate higher-order sensory perception have, until now, proven to be unsuccessful. For example, attempting to alter perception using mental imagery, a cognitive process which utilises similar neural substrates to perception(O'Craven and Kanwisher, 2000), does not increase the vividness of the imagery. Most importantly, mental imagery training has no effect on perception linked to the imagery strategy used during training, as demonstrated with binocular rivalry (BR) between images specifically associated with the mental imagery training(Rademaker and Pearson, 2012). BR is a unique perceptual phenomenon that has been used to provide a window into the unconscious and conscious processes underlying visual perception. It is produced by simultaneously presenting conflicting monocular stimuli to each eye. Paradoxically, the brain cannot form a stable image. Instead, each image randomly competes for exclusive perceptual dominance. Until now, producing unconscious shifts in higher-order perception by directly modifying brain function has proven to be unsuccessful.

Neurofeedback training using realtime functional magnetic resonance imaging (rt-fMRI) is an emerging technique which allows participants to control target brain regions by voluntarily modulating online feedback of activity in those regions(Sitaram et al., 2016). Feedback is typically provided via a visual interface during concurrent MR scanning. Online modulation of the Blood Oxygen Level-Dependent (BOLD) signal using neurofeedback involves abstract cognitive strategies, as well as mental imagery that maybe explicitly linked to the brain region-of-interest (ROI). This approach can produce changes in behaviour through the functional modulation of trained brain regions, including low-order visual perception (e.g. grating orientation, colour) by modulating primary retinotopic cortex(Amano et al., 2016; Shibata et al., 2011), pain and craving by modulating anterior cingulate cortex(deCharms et al., 2004; Li et al., 2013), and motor function by modulating supplementary motor area and primary motor cortex(Blefari et al., 2015; Subramanian et al., 2011). We hypothesised that rt-fMRI neurofeedback might prove more powerful than previous approaches, such as mental imagery alone, in enabling participants to modify brain activity associated with higher-order visual perception, and consequently directly influence how they perceive the world.

To test this hypothesis, we trained human participants using mental imagery *combined* with neurofeedback to voluntarily control the difference in activation between two higher-order visual cortical regions (Fusiform Face Area, FFA and Parahippocampal Place area, PPA). The human FFA responds strongly to faces(Kanwisher et al., 1997; McCarthy et al., 1997), but not to other types of non-face stimuli, while the PPA responds to houses and places, but not faces(Epstein and Kanwisher, 1998). Further, both of these regions activate during mental imagery of faces or places respectively, even in the absence of visual stimuli(O’Craven and Kanwisher, 2000). The differential response properties of these two regions enabled participants in the study to have a visually presented neurofeedback training signal that represented the difference in activation between the two regions i.e. a *differential signal*.

The use of a differential signal provided an internal control for global brain activation, and helped focus the training effect on the two selected brain regions in a manner that might not occur with mental imagery training only. We tested participants with a BR task, where they were exposed to rivalrous monocular face and house images, before and after neurofeedback training. During BR, participants are consciously aware of only one of the perceptual stimuli at a time, while the other stimulus is temporarily suppressed. The perceptual fluctuation is spontaneous and stochastic, with both top-down (i.e. cognitive modulation) and bottom-up (i.e. salience-based) processes being implicated(Dayan, 1998; Parker and Alais, 2007; Tong et al., 2006). In this study, the ensuing BR, where perception alternated spontaneously between each monocular view, provided a test of whether neurofeedback training had altered the likelihood of either stimuli entering awareness. We investigated whether any perceptual changes were associated with differences in brain activity and structure (see also Supplementary Materials).

To anticipate our findings, following neurofeedback training, there was a sustained influence on the perceptual dynamics of BR, suggesting functional plasticity. This effect was additionally observed when participants performed concurrent modulation of brain activity during BR. Further, a multivariate analysis of changes in brain structure produced by neurofeedback training predicted changes in BR dynamics.

Methods

Main experiment:

Participants

Ten neurologically normal adult volunteers (24–35 years of age; mean age 28 years; 8 females) with normal or corrected-to-normal visual acuity participated in the experiment. Each participant gave written informed consent. The study was approved by the local ethics committee (UCL Ethics Committee code: 09/H0716/14).

Stimuli and Materials

All visual stimuli were generated and displayed via scripts in MATLAB created with the Cogent 2000 toolbox (http://www.vislab.ucl.ac.uk/cogent_2000.php), on a viewing screen with a visual angle of 23 degrees by 17 degrees, (30 x 26 LCD projector (LT158; NEC). The mirror-mounted viewing screen was set on the top of the scanner bore (optical distance 52 cm). During the neurofeedback sessions, participants saw a fluctuating thermometer bar at the centre of the screen. During the BR sessions only, participants wore a pair of prism glasses. Additionally, a black cardboard divider was placed between the forehead and the screen to ensure that each eye could see one side of the screen only, and provide a stable base for fixation. Two identical box stimuli were displayed side-by-side on the monitor, each with a central white fixation cross (0.68 visual angle) and tile frame surround (11.78 visual angle), upon a uniform grey background (background luminance ¼ 65 Cd/m^2). Optimal perceptual fusion of the two box stimulus images was confirmed with the participant prior to commencing each BR session. Face or house stimuli were presented (20 exemplars each). Responses for durations were obtained via a pair of custom-built, MR-compatible, response boxes.

FMRI Scanning

Scanning was performed on a 3T Allegra head-only scanner (standard transmit-receive head coil). Functional data was acquired with a single-shot gradient echo planar imaging sequence (matrix size, 64x64; field of view, 192x192mm; isotropic in-plane resolution, 3x3 mm; 32 slices with ascending acquisition; slice thickness, 2 mm; slice gap, 1 mm; echo time (TE), 30 ms; repetition time (TR), 1920 ms; flip angle, 90°; receiver bandwidth, 3551 Hz/pixel). Although the nominal slice thickness was 2mm, the effective slice profile achieved in practice is typically larger such that the effective slice thickness is closer to 3mm. Allowing a gap additionally minimised any risk of saturation effects upon excitation of the subsequent slice (again due to imperfect slice profiles). This is particularly important in the case of ascending acquisition order, as used here. Ascending acquisition order was chosen to minimise the impact of any participant motion, which again could lead to saturation effects if the motion resulted in any part of the previously excited slice being re-excited in a time shorter than the TR.

Within each scanning session, double-echo fast, low-angle shot sequence

(FLASH) field maps (TE1, 10 ms; TE2, 12.46 ms; resolution, 3 x 3 x 2 mm; slice gap, 1 mm) were acquired and used to correct geometric distortions.

High Resolution Structural Scans

A whole brain high-resolution T1-weighted structural scan was performed before and after training. This was in addition to structural scans performed on each neurofeedback training day. The scan was a 3D-modified, driven equilibrium Fourier transform (MDEFT) scan (1mm isotropic resolution; matrix size, 256x240 mm; field of view, 256x240 mm; 176 sagittal partitions; TE, 2.4 ms; TR, 7.92 ms; inversion time, 910 ms; flip angle, 15°; readout bandwidth, 195 Hz/pixel; spin tagging in the neck with flip angle 160° to avoid flow artifacts for superposition of functional maps(Deichmann et al., 2004)).

Realtime fMRI Set-up for Neurofeedback

Turbo Brain Voyager(Goebel et al., 2006) was used, with custom realtime image export tools programmed in ICE VA25 (Siemens Healthcare)(Weiskopf et al., 2004), and custom MATLAB based scripts. Participants were shown visual representations of BOLD signal changes in brain regions previously identified with a functional localiser scan (i.e. target ROIs). Realtime data preprocessing encompassed 3D motion correction, smoothing, and incremental linear detrending of time series. The ROI time course(s) were extracted from the prescribed ROI masks, averaged and exported. Signal drift, spikes and high frequency noise were further removed in realtime from the exported time courses with custom MATLAB scripts(Koush et al., 2012). The feedback signal (a ‘fluctuating’ thermometer bar) was displayed to the participants with a delay of 2 s from the acquisition of the image.

Binocular Rivalry Set-up and Behavioural data acquisition

Inside the scanner, participants, wearing custom-made prism glasses, were shown two stimuli equidistant from a central viewing screen divider. During the viewing blocks, a face stimulus and a house stimulus were presented in the left and right hemi-fields respectively. The stimuli were pseudorandomised with regards to which eye received the face or house stimuli. Each viewing block (40 s followed by rest 20 s) was performed with a new pair of stimuli from the pool of 20 stimuli. Six blocks were performed per session, for three sessions.

During the BR sessions, participants pressed one of three buttons to record their percept of ‘face’, ‘house’ or ‘mixed’. The participants were instructed to switch as accurately and rapidly as possible between the three possible button presses linked to the three percepts. This was the only instruction given during pre-training BR and post-training BR, which were identical save for being performed either side of neurofeedback training. Additional instructions were given for two further post-training BR conditions (see below,

Day 5: Post-training BR).

Cumulative dominance durations were calculated, which were equal to the total amount of time each monocular stimulus was perceived, and averaged across blocks. The three percepts were then pooled as follows: (1) ***strategy-related percept*** e.g. face percept for the neurofeedback group advised to use face mental imagery (‘Face’ group) or house percept for the neurofeedback group advised to use house mental imagery (‘House’ group) (2) ***strategy-unrelated percept’*** e.g. house percept for the ‘Face’ group, face percept for the ‘House’ group); and (3) ***‘mixed percept’*****.**

Experimental Outline

The experiment was divided into multiple days, with each participant attending five consecutive scanning days (Figure 1). The participants were split into two groups, with five participants in the ‘face’ group and five participants in the ‘house’ training group.

Day 1: Pre-training BR and Localiser

A **Pre-training BR** scan was performed as described above for all participants. They then underwent a **functional localiser** scan to identify FFA and PPA regions (12 minutes, 16 blocks of face stimuli, 16 blocks of house stimuli, and 20 different exemplars per block). Each stimulus was presented for 600 ms (400 ms interstimulus interval). A one-back task was performed (3 targets per block), requiring a button press upon detection of the same stimulus. Two contrasts were used; Houses vs. Faces and Faces vs. Houses. Using the Juelich histological atlas to provide an anatomical landmarks(Eickhoff et al., 2006, 2005), voxel selection for the ROIs were defined along the ventral and lateral surfaces of the temporal lobe in proximity to the fusiform gyrus for FFA, and lateral to the collateral sulcus in the parahippocampal region for PPA respectively.

Day 2-4: Neurofeedback Sessions

Each neurofeedback training day comprised three scanning sessions, each six blocks of 60 s with an ‘upregulate’ period (40 s) followed by ‘rest’ (20 s). During an upregulation period, participants viewed a fluctuating red bar and a fixed horizontal black bar. The latter was placed towards the top of the screen, and the participants were asked to push the red bar above it. Participants were told that the fluctuating red bar was linked to their brain activity, and that they should drive the red bar up to the level of the black bar using a mental imagery strategy. They were advised to maintain the red bar at that level, for as long as possible, during the ‘upregulate’ period. Participants were told that there was a delay related to the training signal (produced by the hemodynamic response function, HRF) of approximately 6-8 s. During rest, participants were instructed to perform a mental arithmetic task (serial subtraction of 7 from 100).

Controlling the Neurofeedback Training Signal

Participants were pseudorandomised into two groups – a ‘Face’ group and a ‘House’ group. Each group was instructed to use mental imagery strategies. They were given examples of what might work (Figure 1), although the participants could use their own interpretation. Specific examples for the house group were ‘think about your house, or a building you are familiar with such as a school or church’, or ‘think about walking down the road looking at buildings’. Specific examples for the face group were ‘think of faces of people you know’, ‘think of celebrity faces’, or ‘think of memorable faces you have seen recently’. Both groups were instructed to pay close attention to the fluctuating red bar, and to find the best way of pushing the bar up for as much and as long as possible. Both groups were instructed to use whatever strategy worked best, including their own, and to vary the strategy to ensure continuous control of the fluctuating red bar.

Each group was unaware of the precise nature of their feedback signal. During neurofeedback training, the fluctuating red bar was driven by brain activity in which the signal from PPA was subtracted from FFA for the ‘Face’ group, and the reverse subtraction (PPA – FFA) for the ‘House’ group. Participants were trained to modulate a *differential* training signal. Therefore, the ‘Face group’ learned to voluntarily increase the difference in BOLD between FFA and PPA. In contrast, the ‘House group’ learned to voluntarily increase the difference in BOLD between PPA and FFA.

For each group there was a *strategy-related ROI* (e.g. FFA for the Face group and a *strategy-unrelated ROI* (e.g. PPA for the Face group, and vice versa for the House group, Figure 2A).

Day 5: Transfer Session

After the final neurofeedback training session, there were two transfer sessions, each comprising six blocks. Each block lasted 60 s and consisted of an ‘upregulate’ period (40 s) followed by ‘rest’ (20 s). During upregulation, participants were required to drive their brain activity ‘up’, using the mental imagery strategies successfully used to drive the bar during neurofeedback training, but now in the absence of a feedback signal.

Day 5: Post-training BR

All participants then performed post-training BR, with the same set-up described for pre-training BR. Three different BR conditions were performed (2 sessions each) pseudorandomised and counterbalanced across all participants: (1) **Post-training BR**. The instruction was identical to the pre-training BR; (2) **Post-training BR with ‘concurrent trained upregulation’.** Both groups were instructed to use their trained mental imagery strategies that had worked best during the training sessions while simultaneously performing BR; and (3) **Post-training BR with ‘concurrent non-trained mental imagery’**. Participants were instructed to use mental imagery related to either houses if in the ‘Face group’, or faces if in the ‘House group’. Mental imagery was to be performed while concurrently performing BR.

Brain Imaging

Functional data was analysed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm). To allow for T1 equilibration the first five images of each session were discarded. Preprocessing involved bias correction, realignment of each EPI to the mean EPI, unwarping, and co-registration of the functional data to the structural image. Normalisation was not performed, as initial analyses were performed in native space. Data was smoothed with a 6 mm FWHM Gaussian kernel and high-pass filtered (128s cut-off) to remove low-frequency noise, while at the same time preserving as many of the spontaneous fMRI fluctuations as possible(Cordes et al., 2001). Session-specific grand mean scaling was applied with no global normalisation.

Offline ROI Analysis: *Fusiform Face Area and Parahippocampal Place Area*

Neurofeedback

BOLD signals across the 9 training sessions (acquired on Days 2-4) were modeled using a GLM, with regressors for each of the 9 sessions. Boxcar functions were created for the six upregulation blocks, convolved with the canonical HRF. Six regressors for movement and a global constant were included. Beta values from the GLM were averaged across all the voxels in the ROI masks (FFA and PPA ROIs based on the functional localiser). Mean percentage signal change (PSC) was then calculated. For each participant, the differential mean PSC between the two ROIs (i.e. strategy-related ROI minus strategy-unrelated ROI) was calculated across sessions. From this, the average mean PSC across participants over the training was calculated.

Transfer Sessions

Two transfer sessions were performed, with participants performing six blocks of upregulation of brain activity as trained, but now in the absence of a neurofeedback signal. In a similar manner to the neurofeedback sessions (see above), the differential mean PSC between the two ROIs (i.e. strategy-related ROI minus strategy-unrelated ROI) was calculated across sessions, and from this, the average mean PSC across participants over transfer was calculated.

Binocular Rivalry

Boxcar functions were created to model the onset of the BR block, convolved with the canonical HRF, for each BR condition. A GLM was performed at the single participant level. Beta values for each of the trained ROIs were averaged for each condition and adjusted for the global brain signal. Mean percentage signal change (PSC) was then calculated.

For inferential statistical analyses, SPSS 21 (IBM Corp. Armonk, USA) was used to perform ANOVAs and follow-up planned paired sample t-tests, which were two-tailed unless otherwise stated.

Control Experiment- Mental Imagery:

Experimental outline

Ten different participants (age range = 22-39 years, mean age 30. years, 8 females) were recruited for a control BR experiment. They viewed a Dell LCD monitor (width: 43.5 cm; resolution: 1600 900; refresh rate: 60 Hz) from a distance of 43 cm (fixed using a chin rest) through a mirror stereoscope. The stereoscope reflected the left and right sides of the screen into the participants' left and right eyes, so that each eye was presented with only one of the two images (house or face). In order to ensure robust fusion of binocular images, prior to the start of BR task, fusion was achieved for each participant by slowly moving two grey squares from the edge toward the centre of the screen. At the beginning of this process the participants would see two squares. By the end of this process the participants would report when they were seeing one square. All testing took place in a darkened room. During the viewing blocks, a face stimulus and a house stimulus were presented in the left and right hemi-fields respectively. The stimuli were pseudorandomised with regards to which eye received the face or house stimuli. Each viewing block (40 s followed by rest 20 s) was performed with a new pair of stimuli from the pool of twenty stimuli. Six blocks were performed per session, for three sessions. Participants were instructed to indicate a perceptual shift only if the whole exemplar was perceived; any combination or 'patchwork' percept regardless of the predominance of the exemplar category was reported as a 'mixed' percept. The participants were instructed to switch as accurately and rapidly as possible between three possible button presses linked to the three perceptual states (face percept, house percept, mixed percept). This resulted in measures of the cumulative duration of the percept throughout the BR measurement period.

BR was performed in this manner prior to and after 3 days of consecutive mental imagery training (see below).

Mental Imagery Training over 3 Days

Participants returned to perform mental imagery training. Participants were pseudorandomised into two equal groups, and were explicitly advised to use mental imagery strategies that involved faces (‘Face group’) or house/places (‘House group’). Mental imagery was undertaken while viewing a LCD monitor screen with a fixed horizontal black bar. They were told to imagine pushing a bar above the fixed black bar, while performing their mental imagery strategies. Each mental imagery training session comprised three sessions, each including six blocks of 60 s with a ‘perform mental imagery’ period (40 s) followed by ‘rest’ (20 s).

Brain Structural Analysis

The structural analysis was performed using Tensor Based Morphometry (TBM), an emerging computational analysis technique(Ceccarelli et al., 2009; Farbota et al., 2012; Li et al., 2009; Wang et al., 2013; Welch et al., 2013), which is better suited to studies with smaller participant samples. TBM enables longitudinal quantitative assessment by identifying regional structural differences from the gradients of the deformation fields that nonlinearly warp each individual image to the template.

For each participant, high-resolution T1 structural images were reoriented placing the anterior commissure at the MNI origin. Longitudinal nonlinear registration(Ashburner and Ridgway, 2012) was performed to align the two time-points (before and after training) to their within-subject average, characterising the relative volumetric expansion or contraction (as the divergence of a velocity field) of each voxel in each time-point with respect to the average. The within-subject average images were then segmented to produce grey and white matter segmentations for each participant(Ashburner and Friston, 2005). These segmentations were nonlinearly aligned to their group-wise average using Dartel (Ashburner, 2007), and the final Dartel average template was affinely registered to MNI space. The resultant between-subject transformations were then used to spatially normalise the divergence maps of the velocity fields, which were finally smoothed with a 6mm FWHM Gaussian kernel.

Divergence measures for each participant were then extracted within spherical ROIs for FFA and PPA (6 mm). The spheres were centered on coordinates that demonstrated the highest functional activity within the localiser ROIs across training. A t-test was then performed to establish if a specific brain region had changed significantly before versus after training.

Canonical Variate Analysis

We used a Canonical Variate Analysis (CVA) to demonstrate that measures of change in brain activation and brain structure following neurofeedback training predicted changes in behavioural measures. Also known as a multivariate analysis of variance, or ManCova (Friston et al., 2014, 1995)**,** CVA enables statistical inferences to be made about associations between the imaging data, and behavioural data that are distributed over variables. It was chosen for analysis of this dataset because it can accommodate statistical dependencies between multivariate predictor variables (behavioural changes) and multivariate outcome variables (functional or structural measures). Neither the behavioural nor imaging data had to be examined in isolation, which had the advantage that distributed changes could be identified, while minimising the multiple comparisons problem. The behavioural changes for each participant was the change in dominance duration of each the three percepts (e.g. *strategy-related* percept, *strategy-unrelated* percept, *mixed* percept) between the pre-training BR condition and post-training BR (Figure S3), and between the pre-training BR condition and post-training BR with concurrent trained up-regulation (Figure S3). As the behavioural and structural measures were taken prior to and immediately after neurofeedback training, the functional measures for each participant were the change in the different signal between the first and the last training run (e.g. run 1 and run 9). The structural measures for each participant were the divergence measures for each ROI, FFA and PPA (6 mm).

The objective of the CVA was to find the linear combination of outcome variables that was best predicted by a linear mixture (contrast) of structural or functional components. The weights of these linear combinations are called canonical vectors. The canonical variates of the outcome and predictor variables are the expression of each canonical vector in each subject. Other quantities generated by the CVA include Bartlett's approximate chi-squared statistic for Wilks' Lambda and its associated significance, or p-value, which test for the significance of a linear mapping or correlation between the canonical variates (in other words, if one or more pairs of canonical variates show a significant statistical dependency).

Results

We first examined the effect of neurofeedback training on behaviour using three comparisons. We compared perceptual dynamics, specifically cumulative dominance durations, performed during BR before and after training. We then examined the effects of learned upregulation on BR by comparing pre-training BR versus post-training BR with concurrent 'learned' upregulation of brain activity. For the final comparison, we examined the non-trained mental imagery on BR, by comparing perceptual dynamics during pre-training BR versus post-training BR with concurrent non-trained mental imagery. The effects of trained mental imagery were additionally examined separately – see Mental Imagery Control Experiment and Figure S1 (Supplementary Materials).

As the durations of the three percepts were dependent on each other, a change in one percept occurred linked to changes in one or both of the other percepts.

Within Condition Comparisons:

Comparison 1. Pre-training BR vs. Post-training BR

Comparing behavioural measures of pre-training BR and post-training BR indicated an effect of training (Figure 3). Paired t-tests revealed a significant reduction in the cumulative dominance durations (i.e. how long a percept type was perceived) for the strategy-unrelated percept (t(9)=2.88,p=0.02), and a significant increase in mixed percept durations (t(9)=2.74,p=0.02), with no significant change in the dominance duration of the strategy-related percept (t(9)=0.46,p=0.66).

Comparison 2. Pre-training BR vs. Post-training BR with Concurrent Trained Upregulation

Paired t-tests revealed a significant reduction in the duration of the strategy-unrelated percept (t(9)=4.76,p=0.001), and a significant increase in the duration of the mixed percept (t(9)=2.68,p=0.03). There was no significant change in the dominance duration of the strategy-related percept (t(9)=0.53,p=0.61) (Figure 3). The changes in BR dynamics were similar to those observed with pre-training BR vs. post-training BR (Comparison 1).

Comparison 3. Pre-training BR vs. Post-training BR with concurrent Non-trained Mental Imagery

Paired t-tests indicated a significant reduction in the duration of the strategy-related percept (t(9)=2.41, p=0.04), and a significant increase in the duration of the mixed percept (t(9)=2.68, p=0.03). There was no significant change in the dominance duration of the strategy-unrelated percept (t(9)=0.12, p=1.74).

We further examined differences between conditions.

Between Condition Comparisons:

1. Post-training BR vs. Post-training BR with Concurrent Trained Upregulation (Comparison 1 vs. Comparison 2)

There was a significantly greater reduction in the dominance duration of the strategy-unrelated percept (t(9)=2.40, p=0.04) in Comparison 2 as compared with Comparison 1 (Figure 3B). There were no other significant differences between the two comparisons (strategy-related percept: t(9)=0.95, p=0.37; mixed percept: t(9)=0.90, p=0.39).

2. Post-training BR vs. Post-training BR with Non-trained Mental Imagery (Comparison 1 vs. Comparison 3)

There was a significantly greater reduction in the dominance duration of the strategy-related percept (t(9)=3.12, p=0.01) (Figure 3B). There was also a significantly greater increase in the dominance duration of the mixed percept (t(9)=2.62, p=0.03). There were no significant changes in the strategy-unrelated percept (t(9)=0.09, p=0.93)

3. Post-training BR with Concurrent Trained Upregulation vs. Post-training BR with Non-trained Mental Imagery (Comparison 2 vs. Comparison 3)

There was a trend towards reduction in the dominance duration of the strategy-related percept (t(9)=2.23, p=0.05) in Comparison 3 as compared with Comparison 2 (Figure 3B). The other two comparisons were not significant (strategy-related percept: t(9)=0.95, p=0.37; mixed percept: t(9)=1.1, p=0.30).

Functional Changes during Neurofeedback Training

To test if neurofeedback training resulted in progressive learning, we examined whether participants demonstrated increased control of the differential feedback signal over the three training days (Figure 2B). A repeated-measures ANOVA with a factor of training day (3 levels; Days 1-3) demonstrated a significant effect (F(2,16)= 3.74, p= 0.047). Post-hoc t-tests demonstrated a significant increase in the differential signal from Day 2 onwards, suggesting a learning effect (Day 1: t(9)=0.88, p= 0.40; Day 2: t(9)=3.27 p=0.001; Day 3: t(8)=2.75, p=0.02).

Functional Changes during Transfer

Following neurofeedback training and prior to BR, voluntary control of brain activation in the absence of neurofeedback was confirmed in a ‘transfer session’. Differential BOLD activation (strategy-related ROI minus strategy-unrelated ROI) pooled across the two transfer sessions, revealed a significant effect (t(9)=2.38, p= 0.04).

Functional Changes during Binocular Rivalry

We examined task-related BOLD signals in the trained ROIs (FFA and PPA) comparing pre-training BR with post-training BR. We observed significant reductions in BOLD signals in both the strategy-related ROI (t(9)= 3.43, p= 0.007) and strategy-unrelated ROI (t(9)= 2.26, p=0.04), when comparing pre-training BR with post-training BR.

Comparing pre-training BR versus post-training BR with concurrent trained upregulation, there was a significant reduction in the activation level of the strategy-unrelated ROI (t(9)= 2.48, p= 0.03). No significant change was noted for the strategy-related ROI (t(9)= 1.41, p=0.19). We performed one-tailed t-tests as we had an *a priori* hypothesis that following neurofeedback training, participants should be able to increase the difference in BOLD activation between the two trained ROIs (Figure 4).

There were no significant changes for pre-training BR versus post-training BR with concurrent non-trained mental imagery (strategy related ROI: t(9)= 0.82, p= 0.44; strategy unrelated ROI: t(9)= 0.83, p= 0.43).

Mental Imagery Control Experiment

There was no evidence of significant changes in the cumulative dominance durations of any of the three percepts (strategy-related percept t(9)= 0.74, p= 0.48; strategy-unrelated percept, t(9)= 1.00, p= 0.34; mixed percept, t(9)= 2.00, p= 0.07).

Between Group Comparisons with 'Mental Imagery' Control Group

We performed an ANOVA with a within-subjects factor of percepts (*strategy-related percept, strategy-unrelated percept, mixed percept*)) and a between-subjects factor of group (Group 1: neurofeedback, Group 2: mental imagery). There was a main effect of percept (F(2,36)= 4.64, p= 0.02). There was no interaction (F(2,36)= 2.65, p= 0.08) between these two factors.

We performed a second ANOVA with a within-subjects factor of percepts (*strategy-related percept, strategy-unrelated percept, mixed percept*)) and a between-subjects factor of group (Group 1: neurofeedback with concurrent-upregulation, Group 2: mental imagery). There was a main effect of percept (F(2,36)=6.68, p=0.003), and an interaction between percept and group (F(2,36)= 5.29, p= 0.01). Follow-up two-sample t-tests looking at changes in durations of the similar percepts showed a significant difference for the strategy-unrelated percept (t(9)= 2.35, p= 0.04), but not for strategy-related percept (t(9)= 1.04, p= 0.32) or the mixed percept, (t(9)= 2.00, p= 0.08).

We performed a further ANOVA with a within-subjects factor of percepts (*strategy-related percept, strategy-unrelated percept, mixed percept*)) and a between-subjects factor of group (Group 1: neurofeedback with concurrent non-trained mental imagery, Group 2: mental imagery). There was a main effect of percept (F(2,36)= 6.70, p= 0.003), and an interaction between percept and group (F(2,36)= 3.63, p= 0.04). Follow-up two-sample t-tests looking at changes in durations of the similar percepts showed a significant difference for the mixed percept (t(9)= 2.79, p= 0.02, but not for strategy-related percept (t(9)= 1.00, p= 0.86) or the strategy-unrelated percept, (t(9)= 0.29, p= 1.14).

Results - Structural

The results of the longitudinal non-rigid registration were used to determine volume changes in the ROIs by calculating the_divergence of the velocity fields. One-sampled t-tests of these values were used to calculate if any significant structural changes had taken place as a result of neurofeedback training. They were not significant for both ROIs i.e. FFA (t(9)= 0.36, p> 0.05), and PPA (t(9)= 0.46, p>0.05),

Results – Canonical Variate analysis

Plots for comparisons of combined measures in: (1) behaviour (dominance durations for the three perceptual reports) and functional (BOLD changes across training in FFA, PPA); and (2) behaviour and structural measures (measure of the volume changes in FFA and PPA following training) are presented in Figure S4, together with Bartlett's approximate chi-squared statistic for Wilks' Lambda and its p-value, for each comparison.

The participant neurofeedback training measures (i.e. differential BOLD brain activation) had a trend to being correlated with changes in BR behavioural dynamics as recorded during BR with concurrent trained upregulation of brain activation (compared with pre-training BR) (chi-squared value = 12.35, p = 0.05). Comparison of changes in the neurofeedback training measures with behavioural changes during 'simple BR' before and after training was non-significant (chi-square value = 11.43, p= 0.07). Significant correlations were noted between structural changes in both ROIs and the change in BR dynamics produced during concurrent trained upregulation of brain activation (chi-squared value = 19.64, p= 0.03). Comparison of structural measures with behavioural measures during ‘simple BR’ before and after training was non-significant (chi-square value = 13.77, p= 0.09).

Of note, the mapping weights obtained for the behavioural measures and the training-related BOLD measures were independent of the mapping weights obtained for the behavioural measures and the structural measures. This is because these multivariate mapping values were specific to the measures used in the comparisons. Finally, the interpretation of the mapping weights in relation to having a positive or negative value did not indicate a positive or negative change in the values (e.g. an increase or decrease in structural measures). Rather they represent a positive (or negative) contribution to the mapping between the multivariate predictor variable and the outcome variables.

Discussion

Participants learned to differentially regulate the amplitude of BOLD activation in two higher-order visual brain regions, FFA and PPA. This was achieved in realtime, through volitional control using neurofeedback training with rt-fMRI. The use of a '*differential*' training signal was implemented by showing the participants a 'thermometer bar' whose size represented the difference in the mean BOLD signal between the two selected brain regions. By doing this, one of the brain regions acted as an internal control for the other, accounting for potential confounds produced by global changes in brain activation in response to effects such as arousal. Furthermore, specific behavioural effects linked with the direction of change of the differential training signal were obtained, providing a comparison of behavioural metrics for the training effect(Thibault et al., 2018). The effect on visual perception was examined with an independent BR task that employed stimuli specifically engaging these stimulus-selective brain regions (face stimuli for FFA, house stimuli for PPA). During BR, moment-to-moment stochastic alternations between two competing visual percepts are observed, while concurrent brain activity can be recorded and potentially manipulated(Blake et al., 2014; Blake and Logothetis, 2002).

In this study, a change in BR perceptual dynamics was observed following neurofeedback training. Perception of the stimulus linked to neurofeedback training was rendered more stable e.g. strategy-related percept, with a reduction in the perception of the *other* stimulus e.g. strategy-unrelated percept. This behavioural change occurred when comparing pre-training BR with post-training BR, and additionally when participants performed post-training BR while concurrently performing learned ‘upregulation’ of brain activity. We compared pre-training ‘BR’ with three post-training BR conditions: ‘post-training BR’, ‘post-training BR with concurrent trained upregulation’, and ‘post-training BR with concurrent non-trained mental imagery’. The first comparison, examining changes during BR before and after neurofeedback training, showed altered BR dynamics; specifically a reduction in the cumulative dominance duration of the strategy-unrelated percept. These findings are important, as they show that neurofeedback training produced a behavioural effect that was: (1) counter-intuitive in that percept durations were not increased in line with the verbally instructed neurofeedback training strategy, which was initially expected. Rather, percept durations *not* linked to the neurofeedback training strategy (e.g. strategy-unrelated percept) were reduced; (2) aligned with a longstanding finding in the field, namely Levelt’s second proposition (discussed below); and (3) indicative of a lack of demand characteristics (see also Mental Imagery Control Experiment).

There was a significant reduction in the levels of activation in both ROIs, comparing pre-training BR versus post-training BR. This linked neuroimaging finding was unexpected, as the prediction from existing literature(Tong et al., 1998) is that BOLD activation levels in extrastriate visual areas will reflect dominance durations. The expected finding might have been that activation levels would be lower in the strategy-unrelated ROI. Our findings instead showed that both regions were affected by neurofeedback training, as we expected given that participants trained on a differential signal involving both ROIs. Both ROIs demonstrated a reduction in activation, which may reflect an increase in neural efficiency as a result of more precise tuning of neural representations(Gimenez et al., 2014; Haler et al., 1992; Heinzel et al., 2014; Vartanian et al., 2013). The exact mechanisms underlying this gain are unknown, particularly in the context of neurofeedback training and thus warrants further study(Poldrack, 2015).

The purpose of the second comparison (‘pre-training BR’ versus ‘post-training BR with concurrent upregulation’) was to examine if there was an effect of concurrent trained modulation of brain activation on BR dynamics that was additive or different to the effect of neurofeedback training alone. We observed a change in BR dynamics that was similar and greater to that observed for pre-training BR vs. post-training BR, in that there was *more of a reduction* in the mean dominance duration of the strategy-unrelated percept. This confirmed that the effect of trained upregulation was directly aligned with the effect of neurofeedback training on BR dynamics. There was a decrease in the level of BOLD activation in the strategy-unrelated ROI only, with no significant change in the strategy-related ROI. Interestingly, these BOLD activation changes were the same as those observed during neurofeedback training (a reduction in activation levels of the strategy-unrelated ROI, Figure 2B). This provides further evidence for a similar mechanism underlying the changes in BR dynamics following training and for those observed with concurrent trained upregulation. The counter-intuitive effect of training and up-regulation (during BR) on the brain activations in the two ROIs (i.e. opposite to an *a priori* instruction and predicted direction of activation changes) is intriguing and worthy of further investigation(Abel et al., 2015; Bueichekú et al., 2016).

The third comparison (‘pre-training BR’ versus ‘post-training BR with non-trained mental imagery’) served to assess the impact of using a differential training signal, which was hypothesised to have an effect on both ROIs in all participants. It additionally helped reveal the role of *non-trained* mental imagery in the context of prior neurofeedback training. No significant change in brain activation in either ROI was observed. However, BR dynamics changed in a similar manner to the other two post-neurofeedback training BR conditions, with a significant reduction in the duration of the percept not linked to the training strategy used during training. This reduction was significant when comparing changes in perceptual dynamics across conditions. These behavioural findings would therefore suggest that neurofeedback training, despite the lack of a statistically significant BOLD effect, produced a more general effect on the neurobiology of the two trained ROIs. The exact nature of this effect may be complex, given that behavioural changes observed for this condition were opposite to the direction of neurofeedback training, but nonetheless sufficient to produce an effect e.g. ‘House’ group participants specifically underwent neurofeedback training with ‘House-based’ mental imagery strategies, and yet they generated changes in BR dynamics simply by using non-trained ‘face’ based mental imagery strategies during the performance of BR. These behavioural findings are different from Rademaker and Pearson’s work, in which using mental imagery training *did not* produce training-related changes in BR dominance duration. Five successive days of mental imagery training had no effect on BR, with no benefit being conferred by expending increased effort during mental imagery generation(Rademaker and Pearson, 2012). On the other hand, Rademaker and Pearson's findings are in keeping with our own mental imagery control experiment, indicating the relevance of neurofeedback training. We conducted a behavioural control experiment in which a different group of participants performed BR before and after three consecutive days of mental imagery training, which was analogous to the neurofeedback training. The training was again explicitly linked to one of the two stimuli used in BR (face mental imagery for a 'Face group', house mental imagery for a 'House group'). However there was no targeted training strategy for the brain, unlike with the neurofeedback-trained groups. No significant changes in dominance durations of any of the three percepts were observed.

Taken together, these results indicate that short-term intensive training over 3 days on a neurofeedback BOLD signal produced by two brain regions, engages and alters the function and biology of *both* regions. This is specifically supported by the shift in perceptual dynamics during BR following neurofeedback training, and the activation changes observed in both ROIs (see Results: Comparison 1). It is further supported more broadly by the behavioural changes observed in all of the post-neurofeedback training BR conditions, which were not observed in the mental imagery control experiment. Habes et al.(Habes et al., 2016) have previously confirmed that although differential regulation of category-specific visual areas can be achieved after a single day of training, a linked change in BR dynamics was not produced. We therefore suggest that in order for mental imagery to produce a change in perception, it must be linked with neurofeedback-led learning, conducted over a period of days. This may be attributable to the interposition of sleep with sequential daily training. Sleep has been directly linked with the offline processing necessary for the consolidation of neuroprosthetic learning(Gulati et al., 2014) and associated behavioural output(Gulati et al., 2017).

Mental imagery may be utilised for perceptual learning of low-level visual features, and to activate stimulus-selective cortical representations(O’Craven and Kanwisher, 2000; Tartaglia et al., 2009). Similarly, rt-fMRI neurofeedback together with implicit operant reinforcement has been used to unconsciously train patterns of activation in primary visual brain regions(Amano et al., 2016; Shibata et al., 2011) to produce perceptual and associative learning of low-level visual features such as colour and orientation. However, to-date neither approach has successfully yielded changes in higher-order visual perception. In this study, we show that coupling explicitly instructed mental imagery with rt-fMRI neurofeedback training of higher-order visual brain regions produces an unconscious and targeted shift in the perceptual processing of visual stimuli. This result is novel and significant in providing evidence for non-invasively manipulating higher-order brain function, potentially at the level of directly strengthening neural representations to alter higher-order perception(Fahle, 2002; Watanabe et al., 2002, 2001). From a mechanistic perspective, an interesting next step might be to test if unconsciously inducing specific patterns of brain activations related to category-specific stimuli will produce linked shifts in perception in a similar manner to that observed in this study(Watanabe et al., 2017). This would provide more direct evidence of modulating neural representations.

The observed behavioural findings may constitute a neural analogue of Levelt's second proposition(Levelt, 1966), as applied to stimulus perception. The original proposition (see Supplementary Discussion) was based on the physical properties of visual stimuli. It was recently modified to indicate that '*increasing the difference in stimulus strength between the two eyes will primarily act to increase the average perceptual dominance duration of the stronger stimulus'*(Brascamp et al., 2015). Our work may provide evidence for a neural reformulation of BR. Participants were trained on a differential signal, rather than specifically training to increase the level of activation in the strategy-related ROI. During training, they appeared to reduce the level of activation in the strategy-unrelated ROI across the three days, while maintaining a fixed level of activation in the strategy-related ROI (Figure 2B). This difference in activation levels as a result of training was maintained when the participants undertook the transfer sessions, an assessment of upregulation in the absence of neurofeedback. The difference in ROI activation levels may have therefore led to a relative difference in the strengths of the neural representations linked to the visual stimulus categories. In keeping with this view, we observed a reduction in the mean dominance duration of the strategy-unrelated percept. This resulted in greater mean dominance durations of the strategy-related percept, corresponding to the ROI with the strengthened neural representation. On the basis of this, we propose a possible neural analogue of the Levelt’s modified second proposition as follows: ‘*increasing the difference in neural representation strengths between the two brain regions linked to the two monocular visual stimuli will primarily act to increase the average perceptual dominance of the percept linked to the stronger neural representation’.* The effect of this would be to produce unconscious perceptual biasing towards the strengthened percept. This mechanism for perceptual ‘shaping’(Lange et al., 2018) may have real-world application in conditions requiring targeted enhancement of perception such as in threat detection(Miranda et al., 2015), or therapeutically to reduce unwanted or aberrant percepts(Taschereau-Dumouchel et al., 2018).

Several mechanisms have been put forward to explain the neural underpinnings of BR. Of note, known influences on visual perception such as priming and cueing have not been shown to produce changes in BR dominance durations (see also Supplementary Discussion). Neurofeedback with rt-fMRI provides the most direct means of testing neuronal function involved in processing visual stimuli. Using a hierarchical model of BR(Dayan, 1998), it may be proposed that neurofeedback training of higher order brain regions strengthens neuronal representations linked to the processing of specific visual stimuli, leading to unconscious perceptual biasing. Preferential processing of strategy-related stimuli would result in decreased dominance durations of the strategy-unrelated stimuli, as was observed here. The effect of neurofeedback on BR may be further considered within a Bayesian framework(Lange et al., 2018). During BR, the dominant percept at any given time is maintained by the highest posterior probability, at the top of the cortical hierarchy. Stimulus representations at lower levels generate error signals that are compared with top-down predictions. The percept is rendered more or less stable in relation to bottom-up inhibition i.e. the lower the error signal, the more stable the percept(Alink et al., 2010; Hohwy et al., 2008; Summerfield and Koechlin, 2008). In keeping with this, BR dynamics were shifted in the direction of the information represented in the trained visual brain regions. Therefore, perception of the stimulus linked to training was rendered more stable, with a simultaneous reduction in the stability of the perception of the other stimulus, leading to a reduction in its mean dominance duration.

The changes in high-level visual perception following neurofeedback training in this study were associated with structural changes in the trained regions (see Supplementary Materials). We used a multivariate analysis technique, Canonical Variate Analysis, which can accommodate multiple measures of behaviour, structure, and function to help determine the overarching effect of neurofeedback training. The change in BR dynamics (i.e. cumulative dominance durations) was linked with measures of structural changes in FFA, and PPA (Figure S3, Supplementary Materials). These preliminary findings in ten participants suggest that neurofeedback training, even over a relatively short period of time (3 days) can alter perception as a result of plasticity in the trained brain regions(Johansen-Berg et al., 2012; Sagi et al., 2012).

In this study, we provide a direct demonstration of the rapid changes in perception and neural plasticity that can be produced by neurofeedback training of higher-order visual areas using rt-fMRI. Imagery-related activation in higher-order visual cortex, such as the ventral visual areas, are related to semantic content, and are more flexible and abstract(Orban et al., 2014) as compared to early visual cortex. Therefore, the use of higher-order visual areas paired with rt-fMRI neurofeedback training may provide the most potent and generalizable means of enacting a change on complex perception. Neural representations that give rise to prior expectations can be directly shifted in the direction of neurofeedback training, even in the presence of pre-existing expectations. This could lead to targeted enhancement of specific responses during discrete tasks as demonstrated here using BR, or in the reduction of aberrant visual perception, such as hallucinations, for therapeutic effect(Lange et al., 2018).

References

Abel, S., Weiller, C., Huber, W., Willmes, K., Specht, K., 2015. Therapy-induced brain reorganization patterns in aphasia. Brain 138, 1097–1112. doi:10.1093/brain/awv022

Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., Muckli, L., 2010. Stimulus predictability reduces responses in primary visual cortex. J. Neurosci. 30, 2960–6. doi:10.1523/JNEUROSCI.3730-10.2010

Amano, K., Shibata, K., Kawato, M., Sasaki, Y., Watanabe, T., 2016. Learning to Associate Orientation with Color in Early Visual Areas by Associative Decoded fMRI Neurofeedback. Curr. Biol. 26, 1861–1866. doi:10.1016/j.cub.2016.05.014

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. Neuroimage 38, 95–113. doi:10.1016/j.neuroimage.2007.07.007

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839–51. doi:10.1016/j.neuroimage.2005.02.018

Ashburner, J., Ridgway, G.R., 2012. Symmetric diffeomorphic modeling of longitudinal structural MRI. Front. Neurosci. 6, 197. doi:10.3389/fnins.2012.00197

Blake, R., Brascamp, J., Heeger, D.J., Brascamp, J., 2014. Can binocular rivalry reveal neural correlates of consciousness ?

Blake, R., Logothetis, N.K., 2002. Visual competition. Nat. Rev. Neurosci. 3, 13–21. doi:10.1038/nrn701

Blefari, M.L., Sulzer, J., Hepp-Reymond, M.-C., Kollias, S., Gassert, R., 2015. Improvement in precision grip force control with self-modulation of primary motor cortex during motor imagery. Front. Behav. Neurosci. 9, 18. doi:10.3389/fnbeh.2015.00018

Brascamp, J.W., Klink, P.C., Levelt, W.J.M., 2015. The ‘laws’ of binocular rivalry: 50 years of Levelt’s propositions. Vision Res. 109, 20–37. doi:10.1016/j.visres.2015.02.019

Bueichekú, E., Miró-Padilla, A., Palomar-García, M.-Á., Ventura-Campos, N., Parcet, M.-A., Barrós-Loscertales, A., Ávila, C., 2016. Reduced posterior parietal cortex activation after training on a visual search task. Neuroimage 135, 204–213. doi:10.1016/j.neuroimage.2016.04.059

Ceccarelli, A., Rocca, M.A., Pagani, E., Falini, A., Comi, G., Filippi, M., 2009. Cognitive learning is associated with gray matter changes in healthy human individuals: a tensor-based morphometry study. Neuroimage 48, 585–9. doi:10.1016/j.neuroimage.2009.07.009

Cordes, D., Haughton, V.M., Arfanakis, K., Carew, J.D., Turski, P.A., Moritz, C.H., Quigley, M.A., Meyerand, M.E., 2001. Frequencies Contributing to Functional Connectivity in the Cerebral Cortex in “Resting-state” Data. AJNR Am. J. Neuroradiol. 22, 1326–1333.

Dayan, P., 1998. A hierarchical model of binocular rivalry. Neural Comput. 10, 1119–35.

deCharms, R.C., Christoff, K., Glover, G.H., Pauly, J.M., Whitfield, S., Gabrieli, J.D.., 2004. Learned regulation of spatially localized brain activation using real-time fMRI. Neuroimage 21, 436–443. doi:10.1016/j.neuroimage.2003.08.041

Deichmann, R., Schwarzbauer, C., Turner, R., 2004. Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at

1.5 and 3 T. Neuroimage 21, 757–67. doi:10.1016/j.neuroimage.2003.09.062

Eickhoff, S.B., Heim, S., Zilles, K., Amunts, K., 2006. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps, NeuroImage. doi:10.1016/j.neuroimage.2006.04.204

Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data, NeuroImage. doi:10.1016/j.neuroimage.2004.12.034

Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. Nature 392, 598–601. doi:10.1038/33402

Fahle, M., 2002. Perceptual learning: gain without pain? Nat. Neurosci. 5, 923–924. doi:10.1038/nn1002-923

Farbota, K.D.M., Sodhi, A., Bendlin, B.B., McLaren, D.G., Xu, G., Rowley, H. a, Johnson, S.C., 2012. Longitudinal volumetric changes following traumatic brain injury: a tensor-based morphometry study. J. Int. Neuropsychol. Soc. 18, 1006–18. doi:10.1017/S1355617712000835

Friston, K.J., Frith, C.D., Frackowiak, R.S.J., Turner, R., 1995. Characterizing Dynamic Brain Responses with fMRI: A Multivariate Approach. Neuroimage 2, 166–172. doi:10.1006/nimg.1995.1019

Friston, K.J., Kahan, J., Biswal, B., Razi, A., 2014. A DCM for resting state fMRI. Neuroimage 94, 396–407. doi:10.1016/j.neuroimage.2013.12.009

Gimenez, P., Bugescu, N., Black, J.M., Hancock, R., Pugh, K., Nagamine, M., Kutner, E., Mazaika, P., Hendren, R., McCandliss, B.D., Hoeft, F., 2014. Neuroimaging correlates of handwriting quality as children learn to read and write. Front. Hum. Neurosci. 8, 155. doi:10.3389/fnhum.2014.00155

Goebel, R., Esposito, F., Formisano, E., 2006. Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. Hum. Brain Mapp. 27, 392–401. doi:10.1002/hbm.20249

Gulati, T., Guo, L., Ramanathan, D.S., Bodepudi, A., Ganguly, K., 2017. Neural reactivations during sleep determine network credit assignment. Nat. Neurosci. 20, 1277–1284. doi:10.1038/nn.4601

Gulati, T., Ramanathan, D.S., Wong, C.C., Ganguly, K., 2014. Reactivation of emergent task-related ensembles during slow-wave sleep after neuroprosthetic learning. Nat. Neurosci. 17, 1107–13. doi:10.1038/nn.3759

Habes, I., Rushton, S., Johnston, S.J., Sokunbi, M.O., Barawi, K., Brosnan, M., Daly, T., Ihssen, N., Linden, D.E.J., 2016. fMRI neurofeedback of higher visual areas and perceptual biases. Neuropsychologia 85, 208–215. doi:10.1016/j.neuropsychologia.2016.03.031

Haler, R.J., Siegel, B., Tang, C., Abel, L., Buchsbaum, M.S., 1992. Intelligence and Changes in Regional Cerebral Glucose Metabolic Rate Following Learning. Intelligence 16, 415–426.

Heinzel, S., Lorenz, R.C., Brockhaus, W.-R., Wustenberg, T., Kathmann, N., Heinz, A., Rapp, M.A., 2014. Working Memory Load-Dependent Brain Response Predicts Behavioral Training Gains in Older Adults. J. Neurosci. 34, 1224–1233. doi:10.1523/JNEUROSCI.2463-13.2014

Hohwy, J., Roepstorff, A., Friston, K., 2008. Predictive coding explains binocular rivalry: an epistemological review. Cognition 108, 687–701. doi:10.1016/j.cognition.2008.05.010

Johansen-Berg, H., Baptista, C.S., Thomas, A.G., 2012. Human structural plasticity at record speed. Neuron 73, 1058–60. doi:10.1016/j.neuron.2012.03.001

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–11.

Koush, Y., Zvyagintsev, M., Dyck, M., Mathiak, K. a, Mathiak, K., 2012. Signal quality and Bayesian signal processing in neurofeedback based on real-time fMRI. Neuroimage 59, 478–89. doi:10.1016/j.neuroimage.2011.07.076

Lange, F.P. De, Heilbron, M., Kok, P., 2018. How Do Expectations Shape Perception ? Trends Cogn. Sci. 22, 764–779. doi:10.1016/j.tics.2018.06.002

Levelt, W.J.M., 1966. The alternation process in binocular rivalry. Br. J. Psychol. 57, 225–238.

Li, W., He, H., Lu, J., Lv, B., Li, M., Jin, Z., 2009. <title>Detection of whole-brain abnormalities in temporal lobe epilepsy using tensor-based morphometry with DARTEL</title> 7497, 749723-749723–6. doi:10.1117/12.833128

Li, X., Hartwell, K.J., Borckardt, J., Prisciandaro, J.J., Saladin, M.E., Morgan, P.S., Johnson, K.A., LeMatty, T., Brady, K.T., George, M.S., 2013. Volitional reduction of anterior cingulate cortex activity produces decreased cue craving in smoking cessation: a preliminary real-time fMRI study. Addict. Biol. 18, 739–748. doi:10.1111/j.1369-1600.2012.00449.x

McCarthy, G., Puce, A., Gore, J.C., Allison, T., 1997. Face-Specific Processing in the Human Fusiform Gyrus. J. Cogn. Neurosci. 9, 605–610. doi:10.1162/jocn.1997.9.5.605

Miranda, R.A., Casebeer, W.D., Hein, A.M., Judy, J.W., Krotkov, E.P., Laabs, T.L., Manzo, J.E., Pankratz, K.G., Pratt, G.A., Sanchez, J.C., Weber, D.J., Wheeler, T.L., Ling, G.S.F., 2015. DARPA-funded efforts in the development of novel brain – computer interface technologies. J. Neurosci. Methods 244, 52–67.

O’Craven, K.M., Kanwisher, N., 2000. Mental imagery of faces and places activates corresponding stiimulus-specific brain regions. J. Cogn. Neurosci. 12, 1013–23.

Orban, G. a, Zhu, Q., Vanduffel, W., 2014. The transition in the ventral stream from feature to real-world entity representations. Front. Psychol. 5, 695. doi:10.3389/fpsyg.2014.00695

Parker, A., Alais, D., 2007. A bias for looming stimuli to predominate in binocular rivalry. Vision Res. 47, 2661–74. doi:10.1016/j.visres.2007.06.019

Poldrack, R.A., 2015. Is “efficiency” a useful concept in cognitive neuroscience? Dev. Cogn. Neurosci. 11, 12–17. doi:10.1016/j.dcn.2014.06.001

Rademaker, R.L., Pearson, J., 2012. Training Visual Imagery: Improvements of Metacognition, but not Imagery Strength. Front. Psychol. 3, 224. doi:10.3389/fpsyg.2012.00224

Sagi, Y., Tavor, I., Hofstetter, S., Tzur-Moryosef, S., Blumenfeld-Katzir, T., Assaf, Y., 2012. Learning in the Fast Lane: New Insights into Neuroplasticity. Neuron 73, 1195–1203. doi:10.1016/j.neuron.2012.01.025

Shibata, K., Watanabe, T., Sasaki, Y., Kawato, M., 2011. Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. Science 334, 1413–5. doi:10.1126/science.1212003

Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., Weiskopf, N., Blefari, M.L., Rana, M., Oblak, E., Birbaumer, N., Sulzer, J., 2016. Closed-loop brain training: the science of neurofeedback. Nat. Rev. Neurosci. doi:10.1038/nrn.2016.164

Subramanian, L., Hindle, J. V, Johnston, S., Roberts, M. V, Husain, M., Goebel, R., Linden, D., 2011. Real-time functional magnetic resonance imaging neurofeedback for treatment of Parkinson’s disease. J. Neurosci. 31, 16309–17. doi:10.1523/JNEUROSCI.3498-11.2011

Summerfield, C., Koechlin, E., 2008. A neural representation of prior information during perceptual inference. Neuron 59, 336–47. doi:10.1016/j.neuron.2008.05.021

Tartaglia, E.M., Bamert, L., Mast, F.W., Herzog, M.H., 2009. Human perceptual learning by mental imagery. Curr. Biol. 19, 2081–5. doi:10.1016/j.cub.2009.10.060

Taschereau-Dumouchel, V., Cortese, A., Chiba, T., Knotts, J.D., Kawato, M., Lau, H., 2018. Towards an unconscious neural reinforcement intervention for common fears. Proc. Natl. Acad. Sci. 201721572. doi:10.1073/pnas.1721572115

Thibault, R.T., MacPherson, A., Lifshitz, M., Roth, R.R., Raz, A., 2018.

Neurofeedback with fMRI: A critical systematic review. Neuroimage 172, 786–807. doi:10.1016/j.neuroimage.2017.12.071

Tong, F., Meng, M., Blake, R., 2006. Neural bases of binocular rivalry. Trends Cogn. Sci. 10, 502–511. doi:10.1016/j.tics.2006.09.003

Tong, F., Nakayama, K., Vaughan, J.T., Kanwisher, N., 1998. Binocular rivalry and visual awareness in human extrastriate cortex. Neuron 21, 753–9.

Vartanian, O., Jobidon, M.-E., Bouak, F., Nakashima, A., Smith, I., Lam, Q., Cheung, B., 2013. Working memory training is associated with lower prefrontal cortex activation in a divergent thinking task. Neuroscience 236, 186–194. doi:10.1016/j.neuroscience.2012.12.060

Wang, Y., Yuan, L., Shi, J., Greve, A., Ye, J., Toga, A.W., Reiss, A.L., Thompson, P.M., 2013. Applying tensor-based morphometry to parametric surfaces can improve MRI-based disease diagnosis. Neuroimage 74, 209–30. doi:10.1016/j.neuroimage.2013.02.011

Watanabe, T., Náñez, J.E., Koyama, S., Mukai, I., Liederman, J., Sasaki, Y., 2002. Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. Nat. Neurosci. 5, 1003–1009. doi:10.1038/nn915

Watanabe, T., Náñez, J.E., Sasaki, Y., 2001. Perceptual learning without perception. Nature 413, 844–848. doi:10.1038/35101601

Watanabe, T., Sasaki, Y., Shibata, K., Kawato, M., 2017. Advances in fMRI Real-Time Neurofeedback. Trends Cogn. Sci. 21, 997–1010. doi:10.1016/j.tics.2017.09.010

Weiskopf, N., Mathiak, K., Bock, S.W., Scharnowski, F., Veit, R., Grodd, W., Goebel, R., Birbaumer, N., 2004. Principles of a brain-computer interface

(BCI) based on real-time functional magnetic resonance imaging (fMRI). IEEE Trans. Biomed. Eng. 51, 966–70. doi:10.1109/TBME.2004.827063

Welch, K. a, Moorhead, T.W., McIntosh, a M., Owens, D.G.C., Johnstone, E.C., Lawrie, S.M., 2013. Tensor-based morphometry of cannabis use on brain structure in individuals at elevated genetic risk of schizophrenia. Psychol. Med. 43, 2087–96. doi:10.1017/S0033291712002668

Acknowledgements

JE's contribution was supported by a Wellcome Trust Clinical Research Fellowship.

G.R.R and A.R were supported by The Wellcome Trust Centre for Neuroimaging core funding from the Wellcome Trust 091593/Z/10/Z.

EF was supported by core funding of the School of Psychology and Language Sciences, University of Reading.

FS and Y.K were supported by a grant from the Swiss National Science Foundation (BSSG10_155915)

NW's contribution was supported by the BRAINTRAIN European research network (Collaborative Project) supported by the European Commission under the Health Cooperation Work Programme of the 7th Framework Programme (Grant agreement n° 602186). The Wellcome Trust Centre for Neuroimaging is supported by core funding from the Wellcome Trust 0915/Z/10/Z. The Wellcome Trust Centre for Neuroimaging has an institutional research agreement with and receives support from Siemens Healthcare.

GR's contribution was supported by a Wellcome Trust Senior Research Fellowship.

Author contributions

J.E, F.S, and G.R designed the experiment. J.E, J.S.W, G.R.R, F.S and G.R discussed and planned data analysis. J.E collected and analysed the data.

J.E, Y.K, G.R.R and F.S. developed the real-time acquisition and analysis tools. J.E and E.F performed the mental imagery control experiment. A.R performed the CVA analysis. J.E and G.R wrote the paper. All authors read and commented on the manuscript**.**

Figure legends

Figure 1. Experiment procedure schematic.

Stage 1 *Pre-training BR*

Stage 2 *Neurofeedback training:* 10 participants were separated into two groups, a ‘face’ group and a ‘house group’, and were trained to increase a fluctuating thermometer bar (blue bar), up to a fixed mark (orange bar). After the neurofeedback training sessions, the participants performed a transfer session with brain modulation in the absence of neurofeedback signal.

Stage 3 *Post-training BR*: Three types of sessions: a) BR; b) BR with ‘concurrent trained upregulation’; and c) BR with ‘concurrent non-trained mental imagery’.

Figure 2A. Schematic showing group ROIs (FFA and PPA statistical masks) on inflated canonical brains. Activation was extracted from these regions for production of the differential signal for neurofeedback training. The direction of regulation of these ROIs was specific for each group i.e. House Group, PPA up/ FFA down, Face Group, FFA up/ PPA down.

Figure 2B. Mean BOLD signal changes across groups, in the strategy-related ROI (red) and the strategy-unrelated ROI (blue), for each of the nine training sessions. The green line shows the difference in mean BOLD activation between the two brain regions and corresponds to the neurofeedback training signal that participants visualised in the scanner as a fluctuating bar. Error bars show ±1SEM.

Figure 3A. Cumulative dominance durations across participants for pre-training BR, and the three post-training BR sessions: Post-training, Post-training BR with concurrent trained upregulation, and Post-training BR with concurrent non-trained mental imagery. Error bars show ±1SEM. The total duration of each BR block was 40s.

Figure 3B. Changes in cumulative dominance durations for binocular rivalry (BR) sessions, showing comparisons before and after neurofeedback training collapsed across both groups. Error bars indicate ±1SEM

A. Pre/post training BR comparison
B. Pre/post-training BR with concurrent training upregulation
C. Pre/post-training BR with concurrent non-trained mental imagery

***p<0.05. Double **p<0.01. Horizontal brackets indicate significant differences in the changes of cumulative dominance durations (p<0.05) ~ over a bracket indicates p=0.07.**

Figure 4. BOLD activation changes in the trained ROIs, during binocular rivalry (BR) sessions, before and after neurofeedback training. There was a significant reduction in activation in both the strategy-related ROI and the strategy-unrelated ROI following training. When BR was performed with concurrent trained up-regulation, there was a significant further decrease in BOLD activation in the strategy-unrelated ROI only. Error bars indicate ±1SEM. (* $p<0.05$).

Figures

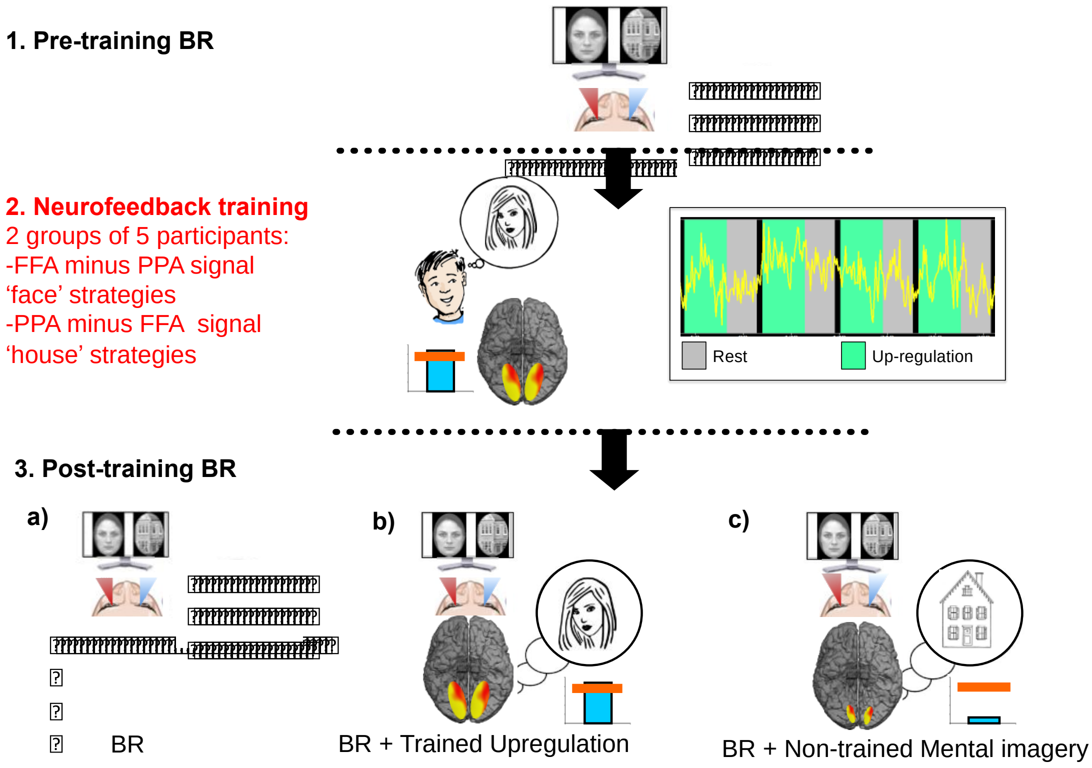

Figure 1

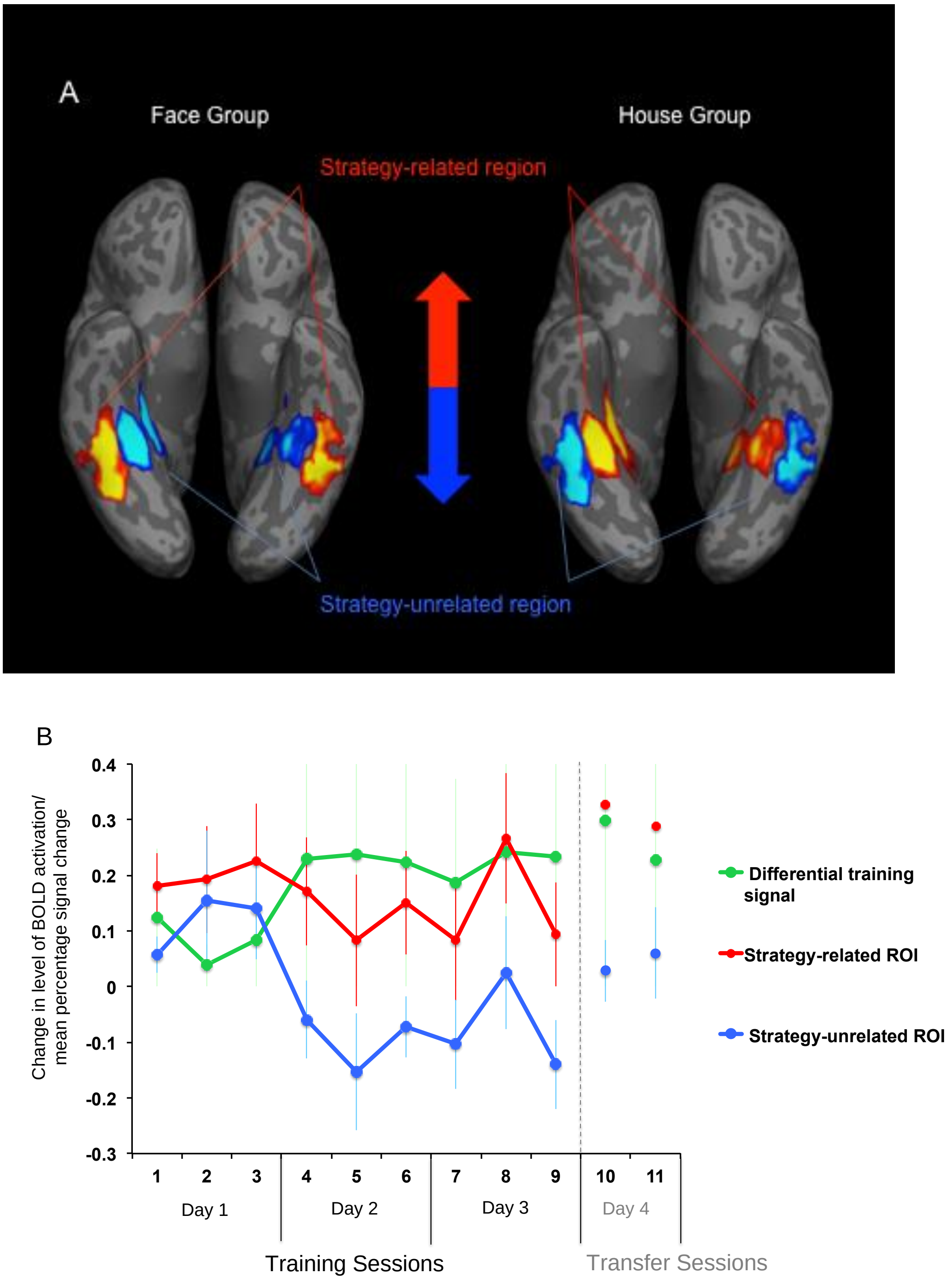

Figure 2A and B.

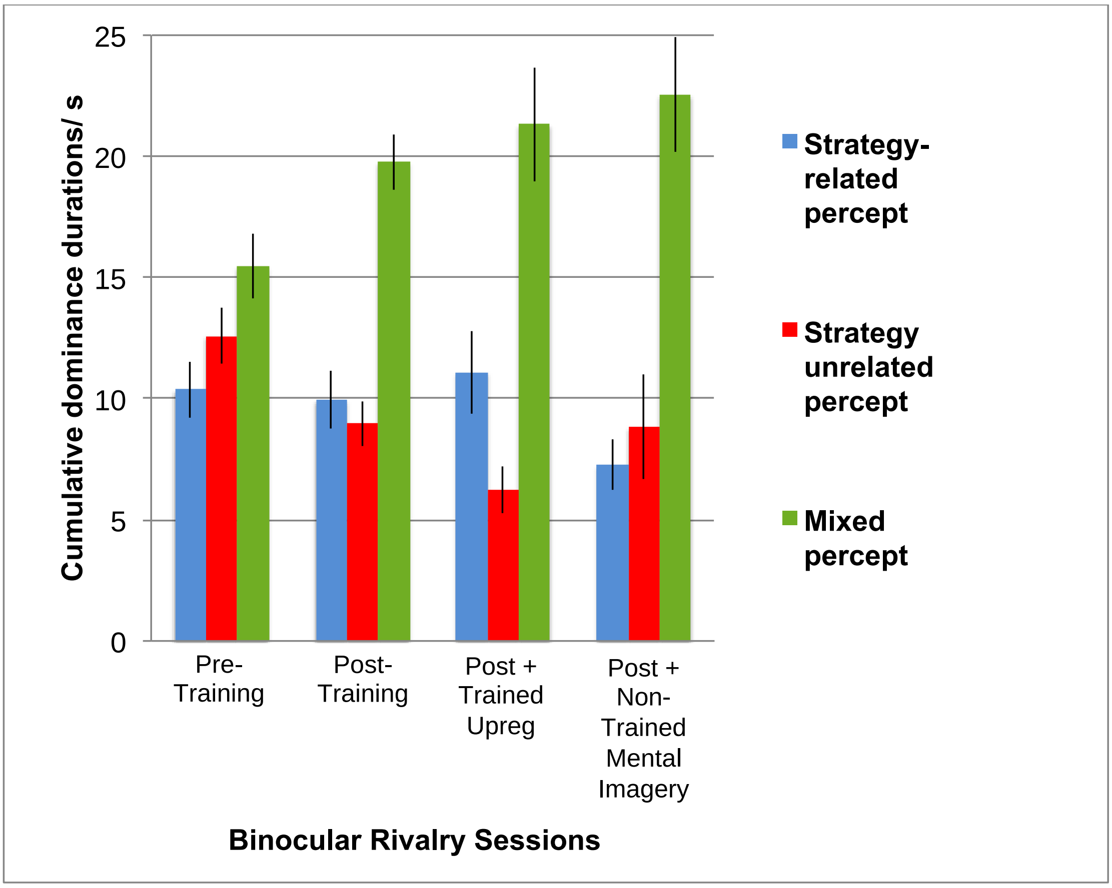
25
20
15
10
5
0
Cumulative dominance durations/ s
Pre-Training
Post-Training
Post + Trained Upreg
Post + Non-Trained Mental Imagery
Binocular Rivalry Sessions
Strategy-related percept
Strategy unrelated percept
Mixed percept

Figure 3A.

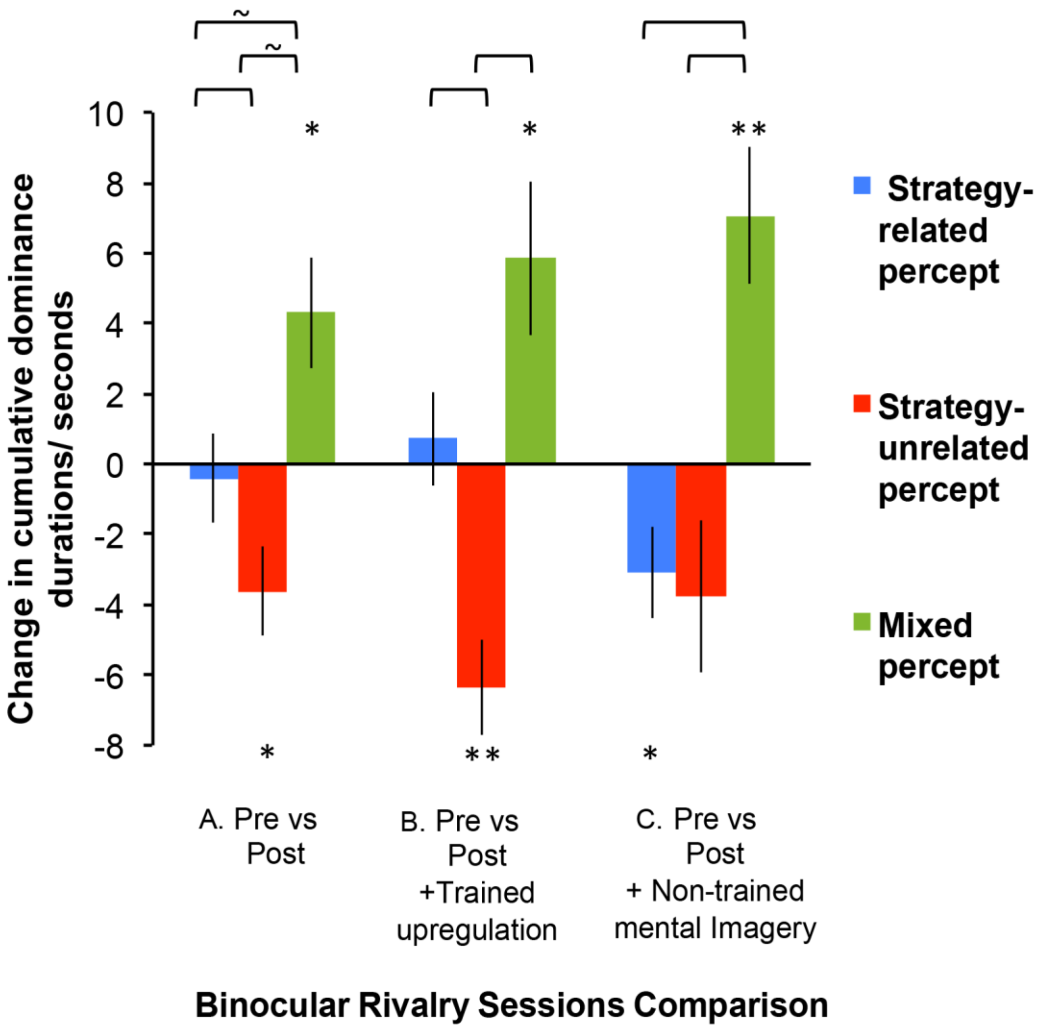
~
~
10
8
6
4
2
0
-2
-4
-6
-8
*
*
**
*
**
*
Change in cumulative dominance durations/ seconds
🟦 Strategy-related percept
🟥 Strategy-unrelated percept
🟩 Mixed percept
A. Pre vs Post
B. Pre vs Post +Trained upregulation
C. Pre vs Post + Non-trained mental Imagery
Binocular Rivalry Sessions Comparison

Figure 3B.

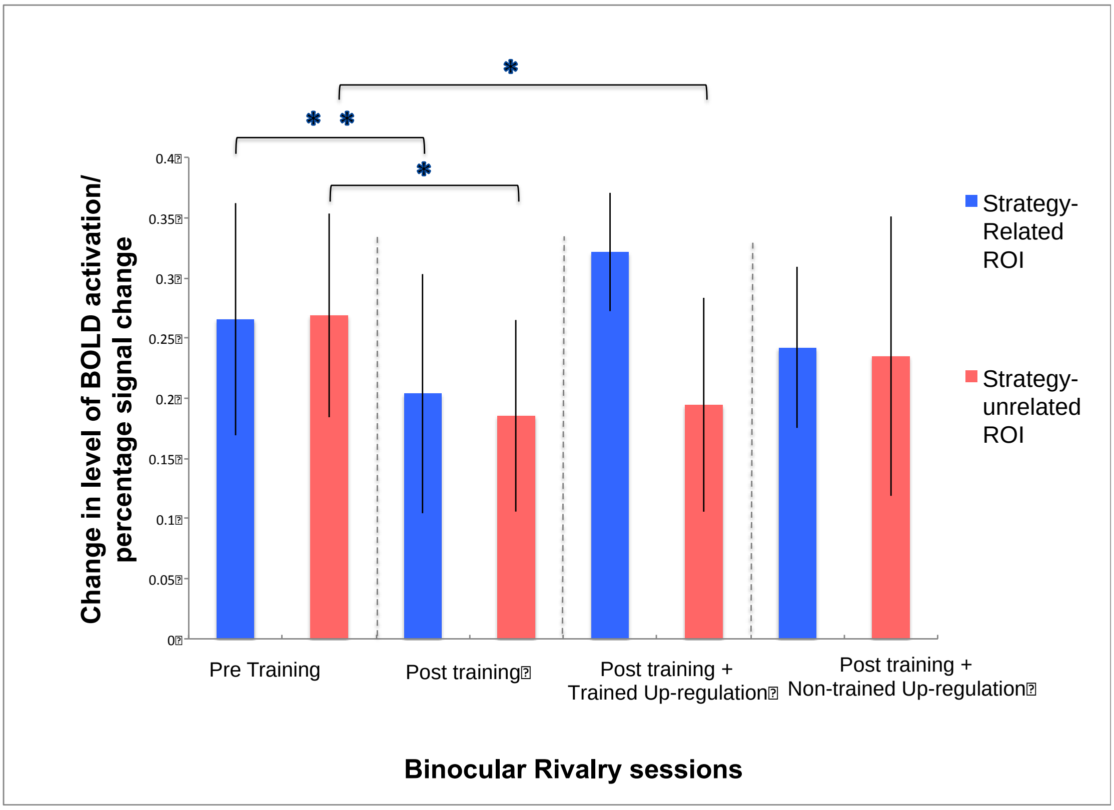
*
* *
*
Change in level of BOLD activation/ percentage signal change
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Strategy-Related ROI
Strategy-unrelated ROI
Pre Training
Post training
Post training + Trained Up-regulation
Post training + Non-trained Up-regulation
Binocular Rivalry sessions

Figure 4.

Supplementary Materials List:

- Supplementary methods
- Supplementary results
- Supplementary discussion
- Supplementary references
- Supplementary figures S1-S4

Supplementary Methods

Learning Effect across Rt-fMRI Neurofeedback Training

The learning effect measures the change in BOLD activation in trained brain region/s across the neurofeedback training sessions. The mean percentage signal change (PSC) for each training run and ROI was calculated and plotted (seeFigure2B in main paper, and compare with S1, S2).

Supplementary Results

Results - Behaviour

Binocular Rivalry – Durations:

See main paper.

Results – Imaging

Strategy-related and Strategy-unrelated ROIs:

We first determined if the differential signal significantly changed over days across participants (see main paper, Result Section). We additionally examined the changes in the two ROIs used to produce the differential signal; the strategy-related ROI and the strategy-unrelated ROI (see Figure 2, main paper).

A one-way ANOVA (with 3 levels corresponding to the 3 training days) revealed a significant reduction in activation in the strategy-unrelated ROI over the 3 days of training ($F(2,16)= 8.71$, $p= 0.003$). On the other hand, a one-way ANOVA for the strategy-related ROI revealed no significant change ($F(2,16)= 0.33$, $p= 0.72$).

Sub-groups:

To assess whether there was any difference between the face and house group during training, an ANOVA was performed on the differential training signal across the 3 training days, with a between-subjects factor with two levels (for the two sub-groups, ‘Face’ and ‘House’). This did not reveal a significant interaction (F(2,14)=0.064, p=0.94) between the two factors.

For neurofeedback training graphs for the two groups (mean percentage signal change over 9 sessions), please see Figures S2 and S3.

Supplementary Discussion

Levelt's Second Proposition, 1966

Levelt's second proposition(Levelt, 1966), as applied to stimulus perception was based on the physical properties of visual stimuli and states: *"Variation of the stimulus strength in one eye will only influence the mean dominance duration of the contralateral eye and not the mean dominance duration of the ipsilateral eye".*

Known Influences on Visual Perception

The role of 'priming' and 'cueing' might also be invoked as possible causes for the perceptual changes observed following neurofeedback training in this study. Prior presentation of a specific orientation grating can cause an increase in the perception of the identical grating during BR. However, dominance durations were unchanged(Denison et al., 2011). Similarly, exogenous cueing prior to BR can increase the probability of the predominant percept being linked to the cue. For example prior to BR, hearing sentences with the word 'face', results in FFA activation(Pelekanos et al., 2011). Nonetheless, no significant change in stimulus dominance between faces and houses on rivalry trials were observed when participants were cued with a word linked to one of the rivalrous stimuli. Dominance durations have also been demonstrated as being immune to the effects of volitional attention(Jung et al., 2016), and reflective of true differences in sensory processing(Dieter et al., 2016). It is therefore unlikely that the perceptual changes produced by neurofeedback training could be ascribed to participant expectation. Evidently, neither altering the level of activity in higher order brain regions involved in perception, nor applying known influences on visual perception, provide a comprehensive explanation for the lasting shifts in perceptual bistability observed following neurofeedback training in this study.

Controlling the Neurofeedback Signal

With regards to the neurofeedback training signal itself (i.e. differential brain activation between two ROIs), there were five potential activation states which could increase the difference between the two brain regions (*strategy-related ROI minus strategy-unrelated ROI*), leading to upregulation of the training signal: These could be: (1) an increase in *strategy-related ROI*; (2) a decrease in *strategy-unrelated ROI*; (3) a combination of the two; (4) a relatively greater increase in *strategy-related ROI* as compared to *strategy-unrelated ROI*; and (5) a relatively greater decrease in the *strategy-unrelated ROI*. Based on our results (Figure 1B in main paper), the mechanism for the upregulation of the differential signal across groups during neurofeedback training appeared to be produced by maintenance of activation in the strategy-related ROI, and a reduction of activation in the strategy-unrelated ROI.

Supplementary figures

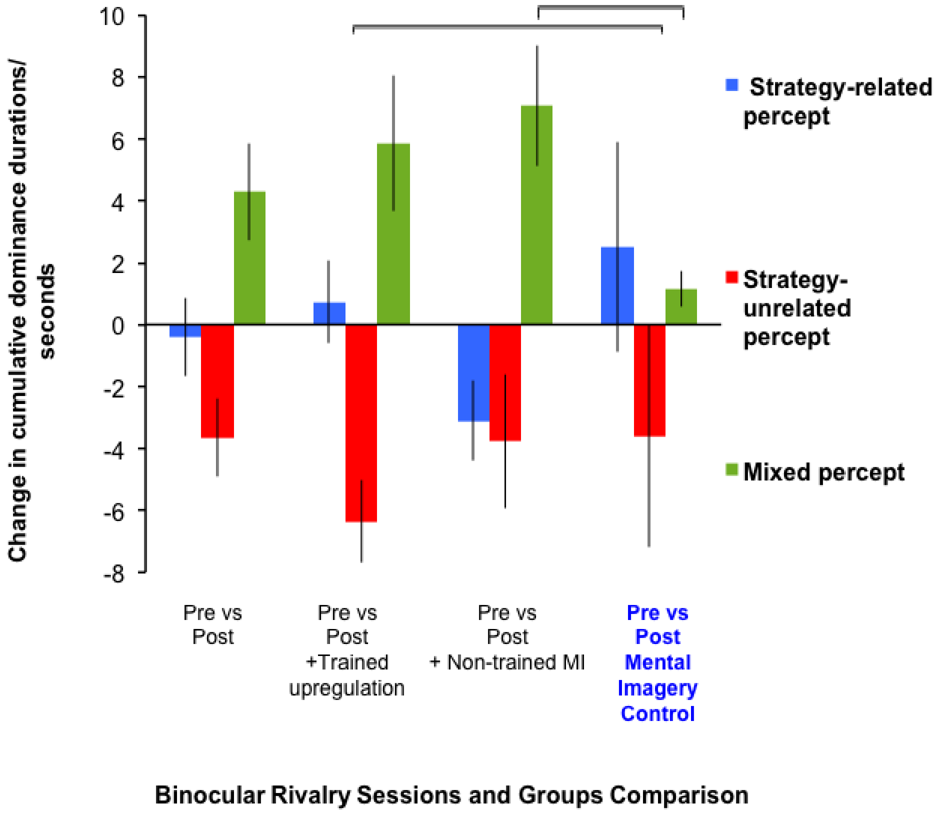

Figure S1. Changes in cumulative dominance durations for binocular rivalry sessions, showing comparisons before and after neurofeedback training. This figure is the analogous to Figure 3B in the main paper, but additionally shows changes in dominance durations for the 'Mental Imagery' control group. Error bars indicate ±1SEM. Horizontal brackets show significant between group comparisons for percepts (p<0.05).

A. Pre vs. Post-training BR comparison

B. Pre vs. Post-training BR with concurrent training up-regulation

C. Pre vs. Post-training BR with concurrent non-trained mental imagery

D. Pre vs. Post training BR comparison for Mental Imagery Control group

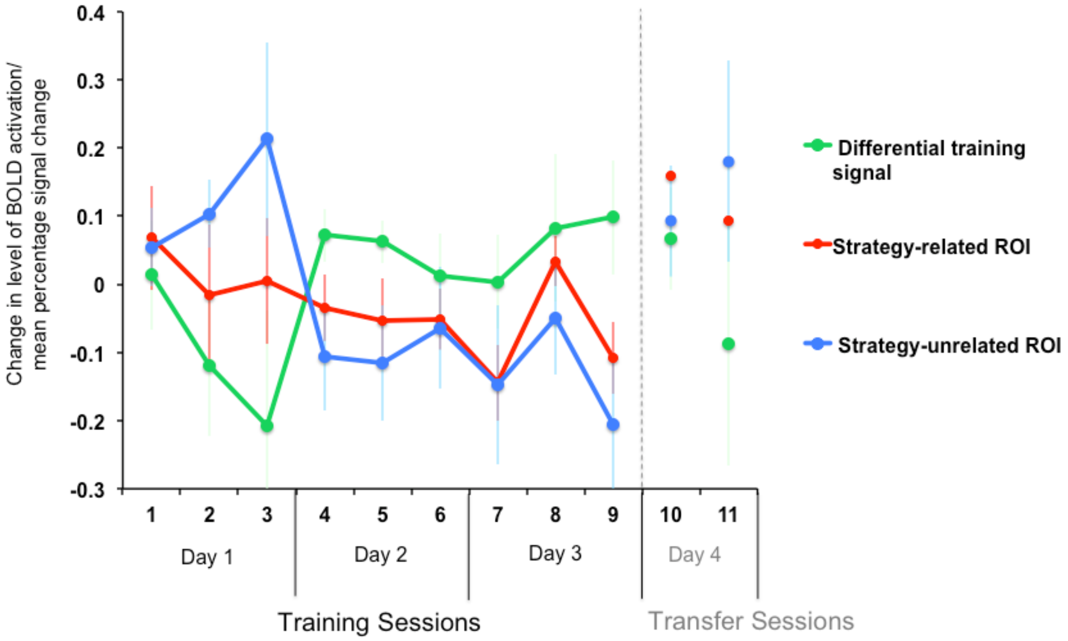

Figure S2. Mean BOLD signal changes across the House group, in the strategy-related brain region (red) and the strategy-unrelated brain region (blue), for each of the nine training sessions. The green line shows the difference in mean BOLD activation between the two brain regions and corresponds to the neurofeedback training signal. Error bars show ±1SEM.

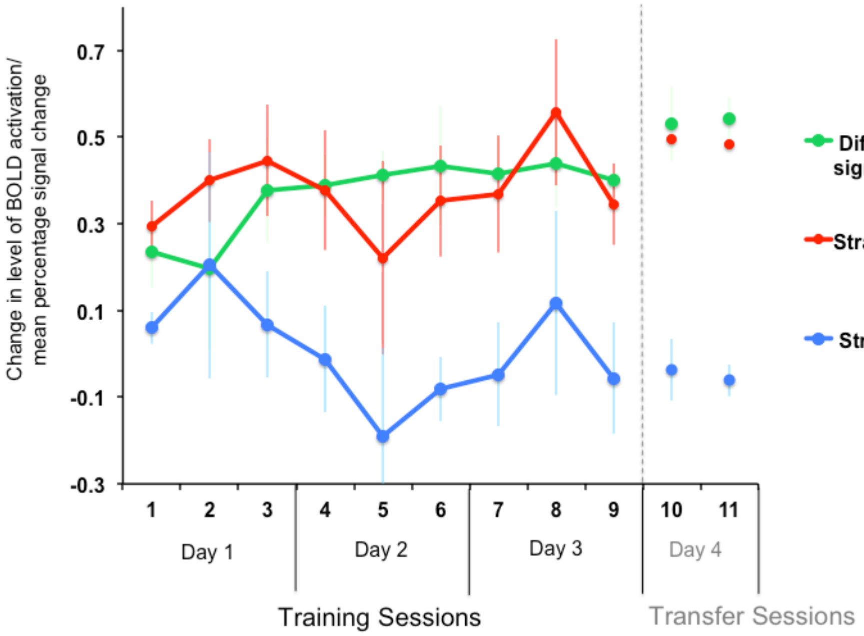

Figure S3. Mean BOLD signal changes across the Face group, in the strategy-related brain region (red) and the strategy-unrelated brain region (blue), for each of the nine training sessions. The green line shows the difference in mean BOLD activation between the two brain regions and corresponds to the neurofeedback training signal. Error bars show ±1SEM.

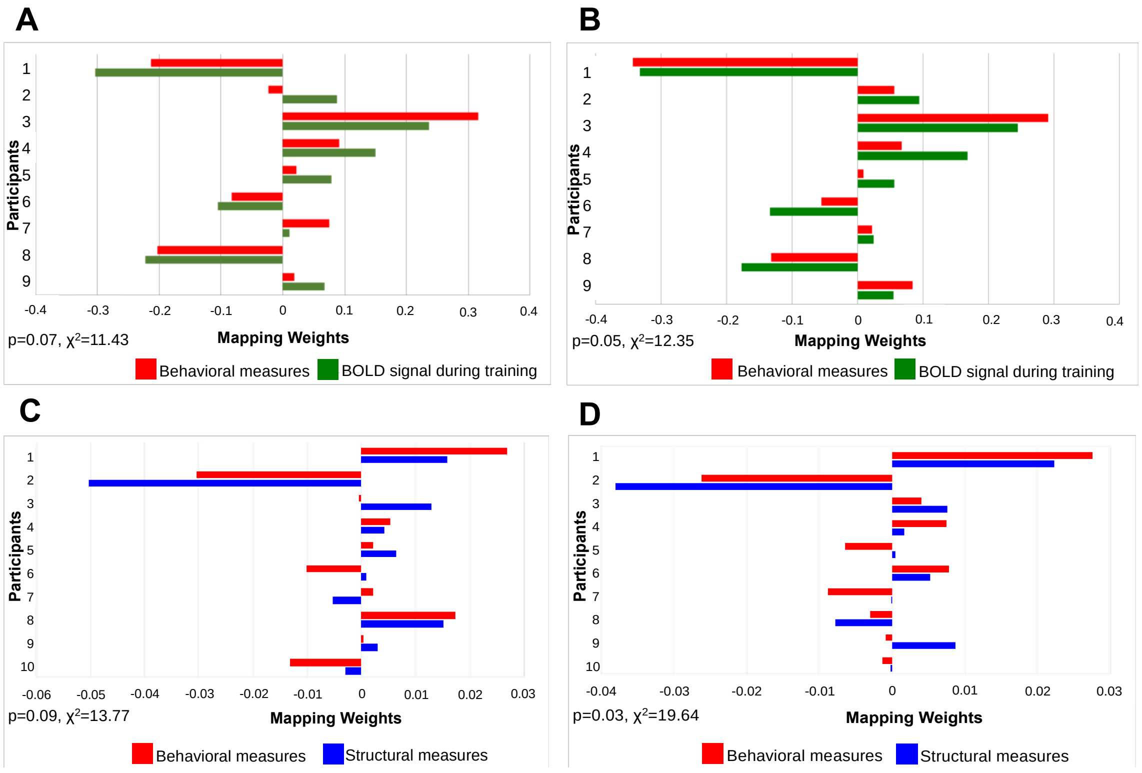

Figure S4. Canonical variate analysis illustrating the correlation between individual behaviour and physiological measures. For each participant mapping weights are shown for pairs of predictor and outcome variables. This approach aims to reveal relationships that may exist between multiple outcome variables following neurofeedback training.

A,B: Comparison of BR behavioural measures (i.e. durations of mixed, strategy-related and strategy-unrelated percepts), and functional BOLD signal changes across training (i.e. differential signal). Nine of the ten participants were included, as one of the participants did not complete all nine training sessions. Participants 1-5 are Face Group, Participants 6-9 are House Group. A shows a non-significant relationship (p= 0.07) between individual participant BR measures (pre vs. post training) and functional BOLD signal changes across training. B shows a non-significant relationship (p= 0.05) between

individual participant BR measures (pre vs. post-training with concurrent trained upregulation) and functional BOLD signal changes across training.

C,D: Comparison of BR behaviour measures (i.e. durations of mixed, strategy-related and strategy-unrelated percepts), and structural measures from FFA and PPA (pre vs. post training). Participants 1-5 are 'Face Group', Participants 6-10 are 'House Group'. C shows a non-significant relationship (p= 0.09) between individual participant BR measures (pre vs. post training) and structural measures from FFA, and PPA (pre vs. post training). D shows a significant relationship (p= 0.03) between individual participant BR measures (pre vs. post-training with concurrent trained upregulation) and structural measures from FFAand PPA (pre vs. post training).

Supplementary References

Denison, R.N., Piazza, E. a, Silver, M. a, 2011. Predictive Context Influences Perceptual Selection during Binocular Rivalry. Front. Hum. Neurosci. 5, 166. doi:10.3389/fnhum.2011.00166

Dieter, K.C., Sy, J.L., Blake, R., 2016. Individual differences in sensory eye dominance reflected in the dynamics of binocular rivalry. Vision Res. doi:10.1016/j.visres.2016.09.014

Jung, Y., Kang, M.-S., Chong, S.C., 2016. Effect of Attention on the Initiation of Binocular Rivalry. Perception 45, 492–504. doi:10.1177/0301006615622324

Levelt, W.J.M., 1966. The alternation process in binocular rivalry. Br. J. Psychol. 57, 225–238.

Pelekanos, V., Roumani, D., Moutoussis, K., 2011. The effects of categorical and linguistic adaptation on binocular rivalry initial dominance. Front. Hum. Neurosci. 5, 187. doi:10.3389/fnhum.2011.00187