Volitional modulation of higher-order visual cortex alters human perception

Article

Accepted Version

www.reading.ac.uk/centaur

CentAUR

1   **TITLE:**

2

3   Volitional Modulation of Higher-order Visual Cortex Alters Human Perception

4

5   **AUTHORS:**

6

7   Jinendra Ekanayake[1,2], Gerard R Ridgway[3], Joel S Winston[1,2], Eva

8   Feredoes[4], Adeel Razi[1,9], Yury Koush[10], Frank Scharnowski[5,6,7], Nikolaus

9   Weiskopf[1,8], Geraint Rees[1,2]

10

11 **AFFILIATIONS:**

12

13 [1]Wellcome Trust Centre for Human Neuroimaging, UCL Institute of Neurology,

14 London, UK

15 [2]UCL Institute of Cognitive Neuroscience, University College London

16 [3]Oxford Centre for Functional MRI of the Brain, University of Oxford, Oxford, UK

17 [4]School of Psychology and Language Sciences, University of Reading, UK

18 [5]Psychiatric University Hospital, University of Zürich, Lenggstrasse 31, 8032 Zürich,

19 Switzerland

20 [6]Neuroscience Center Zürich, University of Zürich and Swiss Federal Institute of

21 Technology, Winterthurerstr. 190, 8057 Zürich, Switzerland

22 [7]Zürich Center for Integrative Human Physiology (ZIHP), University of Zürich,

23 Winterthurerstr. 190, 8057 Zürich, Switzerland

24 [8]Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain

25 Sciences, Leipzig, Germany

26 [9]Department of Electronic Engineering, NED University of Engineering and

27 Technology, Karachi, Pakistan

28   [10]Department of Radiology and Biomedical Imaging, Yale University, 300 Cedar

29   Street, New Haven, CT 06519, USA

30

31   **CORRESPONDING AUTHOR:**

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

33   **ADDRESS:**

34   Institute of Cognitive Neuroscience,

35   12 Queen Square, WC1N 3AR

36

37   **Conflict of Interest: Nil**

38

39 **<u>Abstract</u>**

40

41 Can we change our perception by controlling our brain activation? Awareness

42 during binocular rivalry is shaped by the alternating perception of different

43 stimuli presented separately to each monocular view. We tested the possibility

44 of causally influencing the likelihood of a stimulus entering awareness. To do

45 this, participants were trained with neurofeedback, using realtime functional

46 magnetic resonance imaging (rt-fMRI), to differentially modulate activation in

47 stimulus-selective visual cortex representing each of the monocular images.

48 Neurofeedback training led to altered bistable perception associated with

49 activity changes in the trained regions. The degree to which training

50 influenced perception predicted changes in grey and white matter volumes of

51 these regions. Short-term intensive neurofeedback training therefore sculpted

52 the dynamics of visual awareness, with associated plasticity in the human

53 brain.

54

55 Word count 123 (150 max)

56

57

58

Highlights

59
60

61 • Unconscious biasing of higher-order visual perception with realtime

62 fMRI neurofeedback.

63 • Participants unknowingly modulated two brain regions to control

64 feedback signal

65 • Short-term neurofeedback training over 3 days induced functional

66 plasticity

67 • Neurofeedback may strengthen neural representations and alter prior

68 expectations

69 • Potential avenue for behavioural shaping and therapeutic reduction of

70 aberrant perception

71

72

Introduction

73   **Introduction**

74

75   The ability to causally modify how we perceive the world has potential

76   implications in health and disease. Altering perceptual biases, which may be

77   conscious or unconscious, could modify pathological perception such as

78   hallucinations, or provide a means of selective cognitive

79   enhancement(Miranda et al., 2015). Such attempts to deliberately manipulate

80   higher-order sensory perception have, until now, proven to be unsuccessful.

81   For example, attempting to alter perception using mental imagery, a cognitive

82   process which utilises similar neural substrates to perception(O'Craven and

83   Kanwisher, 2000), does not increase the vividness of the imagery. Most

84   importantly, mental imagery training has no effect on perception linked to the

85   imagery strategy used during training, as demonstrated with binocular rivalry

86   (BR) between images specifically associated with the mental imagery

87   training(Rademaker and Pearson, 2012). BR is a unique perceptual

88   phenomenon that has been used to provide a window into the unconscious

89   and conscious processes underlying visual perception. It is produced by

90   simultaneously presenting conflicting monocular stimuli to each eye.

91   Paradoxically, the brain cannot form a stable image. Instead, each image

92   randomly competes for exclusive perceptual dominance. Until now, producing

93   unconscious shifts in higher-order perception by directly modifying brain

94   function has proven to be unsuccessful.

95

96   Neurofeedback training using realtime functional magnetic resonance imaging

97   (rt-fMRI) is an emerging technique which allows participants to control target

98   brain regions by voluntarily modulating online feedback of activity in those

99   regions(Sitaram et al., 2016). Feedback is typically provided via a visual

100   interface during concurrent MR scanning. Online modulation of the Blood

101   Oxygen Level-Dependent (BOLD) signal using neurofeedback involves

102   abstract cognitive strategies, as well as mental imagery that maybe explicitly

103   linked to the brain region-of-interest (ROI). This approach can produce

104   changes in behaviour through the functional modulation of trained brain

105   regions, including low-order visual perception (e.g. grating orientation, colour)

106   by modulating primary retinotopic cortex(Amano et al., 2016; Shibata et al.,

107   2011), pain and craving by modulating anterior cingulate cortex(deCharms et

108   al., 2004; Li et al., 2013), and motor function by modulating supplementary

109   motor area and primary motor cortex(Blefari et al., 2015; Subramanian et al.,

110   2011). We hypothesised that rt-fMRI neurofeedback might prove more

111   powerful than previous approaches, such as mental imagery alone, in

112   enabling participants to modify brain activity associated with higher-order

113   visual perception, and consequently directly influence how they perceive the

114   world.

115

116   To test this hypothesis, we trained human participants using mental imagery

117   *combined* with neurofeedback to voluntarily control the difference in activation

118   between two higher-order visual cortical regions (Fusiform Face Area, FFA

119   and Parahippocampal Place area, PPA). The human FFA responds strongly

120   to faces(Kanwisher et al., 1997; McCarthy et al., 1997), but not to other types

121   of non-face stimuli, while the PPA responds to houses and places, but not

122   faces(Epstein and Kanwisher, 1998). Further, both of these regions activate

123   during mental imagery of faces or places respectively, even in the absence of

124   visual stimuli(O'Craven and Kanwisher, 2000).   The differential response

125   properties of these two regions enabled participants in the study to have a

126   visually presented neurofeedback training signal that represented the

127   difference in activation between the two regions i.e. a *differential signal*.

128

129   The use of a differential signal provided an internal control for global brain

130   activation, and helped focus the training effect on the two selected brain

131   regions in a manner that might not occur with mental imagery training only.

132   We tested participants with a BR task, where they were exposed to rivalrous

133   monocular face and house images, before and after neurofeedback training.

134   During BR, participants are consciously aware of only one of the perceptual

135   stimuli at a time, while the other stimulus is temporarily suppressed. The

136   perceptual fluctuation is spontaneous and stochastic, with both top-down (i.e.

137   cognitive modulation) and bottom-up (i.e. salience-based) processes being

138   implicated(Dayan, 1998; Parker and Alais, 2007; Tong et al., 2006). In this

139   study, the ensuing BR, where perception alternated spontaneously between

140   each monocular view, provided a test of whether neurofeedback training had

141   altered the likelihood of either stimuli entering awareness. We investigated

142   whether any perceptual changes were associated with differences in brain

143   activity and structure (see also Supplementary Materials).

144

145   To anticipate our findings, following neurofeedback training, there was a

146   sustained influence on the perceptual dynamics of BR, suggesting functional

147   plasticity. This effect was additionally observed when participants performed

148   concurrent modulation of brain activity during BR. Further, a multivariate

149   analysis of changes in brain structure produced by neurofeedback training

150   predicted changes in BR dynamics.

151

152 **Methods**
153
154
155 **Main experiment:**
156

157 **Participants**

158

159 Ten neurologically normal adult volunteers (24–35 years of age; mean age 28

160 years; 8 females) with normal or corrected-to-normal visual acuity participated

161 in the experiment. Each participant gave written informed consent. The study

162 was approved by the local ethics committee (UCL Ethics Committee code:

163 09/H0716/14).

164

165 **Stimuli and Materials**

166

167 All visual stimuli were generated and displayed via scripts in MATLAB created

168 with the Cogent 2000 toolbox (http://www.vislab.ucl.ac.uk/cogent_2000.php),

169 on a viewing screen with a visual angle of 23 degrees by 17 degrees, (30 x 26

170 LCD projector (LT158; NEC). The mirror-mounted viewing screen was set on

171 the top of the scanner bore (optical distance 52 cm). During the

172 neurofeedback sessions, participants saw a fluctuating thermometer bar at

173 the centre of the screen. During the BR sessions only, participants wore a pair

174 of prism glasses. Additionally, a black cardboard divider was placed between

175 the forehead and the screen to ensure that each eye could see one side of

176 the screen only, and provide a stable base for fixation. Two identical box

177 stimuli were displayed side-by-side on the monitor, each with a central white

178 fixation cross (0.68 visual angle) and tile frame surround (11.78 visual angle),

9

179   upon a uniform grey background (background luminance ¼ 65 Cd/m^2).

180   Optimal perceptual fusion of the two box stimulus images was confirmed with

181   the participant prior to commencing each BR session. Face or house stimuli

182   were presented (20 exemplars each). Responses for durations were obtained

183   via a pair of custom-built, MR-compatible, response boxes.

184

185   **FMRI Scanning**

186

187   Scanning was performed on a 3T Allegra head-only scanner (standard

188   transmit-receive head coil). Functional data was acquired with a single-shot

189   gradient echo planar imaging sequence (matrix size, 64x64; field of view,

190   192x192mm; isotropic in-plane resolution, 3x3 mm; 32 slices with ascending

191   acquisition; slice thickness, 2 mm; slice gap, 1 mm; echo time (TE), 30 ms;

192   repetition time (TR), 1920 ms; flip angle, 90°; receiver bandwidth, 3551

193   Hz/pixel). Although the nominal slice thickness was 2mm, the effective slice

194   profile achieved in practice is typically larger such that the effective slice

195   thickness is closer to 3mm. Allowing a gap additionally minimised any risk of

196   saturation effects upon excitation of the subsequent slice (again due to

197   imperfect slice profiles). This is particularly important in the case of ascending

198   acquisition order, as used here. Ascending acquisition order was chosen to

199   minimise the impact of any participant motion, which again could lead to

200   saturation effects if the motion resulted in any part of the previously excited

201   slice being re-excited in a time shorter than the TR.

202

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

204 (FLASH) field maps (TE1, 10 ms; TE2, 12.46 ms; resolution, 3 x 3 x 2 mm;

205 slice gap, 1 mm) were acquired and used to correct geometric distortions.

206

207 **High Resolution Structural Scans**

208

209 A whole brain high-resolution T1-weighted structural scan was performed

210 before and after training. This was in addition to structural scans performed on

211 each neurofeedback training day. The scan was a 3D-modified, driven

212 equilibrium Fourier transform (MDEFT) scan (1mm isotropic resolution; matrix

213 size, 256x240 mm; field of view, 256x240 mm; 176 sagittal partitions; TE, 2.4

214 ms; TR, 7.92 ms; inversion time, 910 ms; flip angle, 15°; readout bandwidth,

215 195 Hz/pixel; spin tagging in the neck with flip angle 160° to avoid flow

216 artifacts for superposition of functional maps(Deichmann et al., 2004)).

217

218 **Realtime fMRI Set-up for Neurofeedback**

219

220 Turbo Brain Voyager(Goebel et al., 2006) was used, with custom realtime

221 image export tools programmed in ICE VA25 (Siemens Healthcare)(Weiskopf

222 et al., 2004), and custom MATLAB based scripts. Participants were shown

223 visual representations of BOLD signal changes in brain regions previously

224 identified with a functional localiser scan (i.e. target ROIs). Realtime data

225 preprocessing encompassed 3D motion correction, smoothing, and

226 incremental linear detrending of time series. The ROI time course(s) were

227 extracted from the prescribed ROI masks, averaged and exported. Signal

228 drift, spikes and high frequency noise were further removed in realtime from

229   the exported time courses with custom MATLAB scripts(Koush et al., 2012).

230   The feedback signal (a 'fluctuating' thermometer bar) was displayed to the

231   participants with a delay of 2 s from the acquisition of the image.

232

233   **Binocular Rivalry Set-up and Behavioural data acquisition**

234

235   Inside the scanner, participants, wearing custom-made prism glasses, were

236   shown two stimuli equidistant from a central viewing screen divider. During

237   the viewing blocks, a face stimulus and a house stimulus were presented in

238   the left and right hemi-fields respectively. The stimuli were pseudorandomised

239   with regards to which eye received the face or house stimuli. Each viewing

240   block (40 s followed by rest 20 s) was performed with a new pair of stimuli

241   from the pool of 20 stimuli. Six blocks were performed per session, for three

242   sessions.

243

244   During the BR sessions, participants pressed one of three buttons to record

245   their percept of 'face', 'house' or 'mixed'. The participants were instructed to

246   switch as accurately and rapidly as possible between the three possible

247   button presses linked to the three percepts. This was the only instruction

248   given during pre-training BR and post-training BR, which were identical save

249   for being performed either side of neurofeedback training. Additional

250   instructions were given for two further post-training BR conditions (see below,

251

252   **Day 5: Post-training BR**).

253

254 Cumulative dominance durations were calculated, which were equal to the

255 total amount of time each monocular stimulus was perceived, and averaged

256 across blocks. The three percepts were then pooled as follows: (1) **strategy-**

257 **related percept** e.g. face percept for the neurofeedback group advised to use

258 face mental imagery ('Face' group) or house percept for the neurofeedback

259 group advised to use house mental imagery ('House' group) (2) **strategy-**

260 **unrelated percept'** e.g. house percept for the 'Face' group, face percept for

261 the 'House' group); and (3) **'mixed percept'.**

262

263 **Experimental Outline**

264

265 The experiment was divided into multiple days, with each participant attending

266 five consecutive scanning days (Figure 1). The participants were split into two

267 groups, with five participants in the 'face' group and five participants in the

268 'house' training group.

269

270 **Day 1: Pre-training BR and Localiser**

271 A **Pre-training BR** scan was performed as described above for all

272 participants. They then underwent a **functional localiser** scan to identify FFA

273 and PPA regions (12 minutes, 16 blocks of face stimuli, 16 blocks of house

274 stimuli, and 20 different exemplars per block). Each stimulus was presented

275 for 600 ms (400 ms interstimulus interval). A one-back task was performed (3

276 targets per block), requiring a button press upon detection of the same

277 stimulus. Two contrasts were used; Houses vs. Faces and Faces vs. Houses.

278 Using the Juelich histological atlas to provide an anatomical

279 landmarks(Eickhoff et al., 2006, 2005), voxel selection for the ROIs were

280 defined along the ventral and lateral surfaces of the temporal lobe in proximity

281 to the fusiform gyrus for FFA, and lateral to the collateral sulcus in the

282 parahippocampal region for PPA respectively.

283

284 **Day 2-4: Neurofeedback Sessions**

285 Each neurofeedback training day comprised three scanning sessions, each

286 six blocks of 60 s with an 'upregulate' period (40 s) followed by 'rest' (20 s).

287 During an upregulation period, participants viewed a fluctuating red bar and a

288 fixed horizontal black bar. The latter was placed towards the top of the screen,

289 and the participants were asked to push the red bar above it. Participants

290 were told that the fluctuating red bar was linked to their brain activity, and that

291 they should drive the red bar up to the level of the black bar using a mental

292 imagery strategy. They were advised to maintain the red bar at that level, for

293 as long as possible, during the 'upregulate' period. Participants were told that

294 there was a delay related to the training signal (produced by the

295 hemodynamic response function, HRF) of approximately 6-8 s. During rest,

296 participants were instructed to perform a mental arithmetic task (serial

297 subtraction of 7 from 100).

298

299 *Controlling the Neurofeedback Training Signal*

300

301 Participants were pseudorandomised into two groups – a 'Face' group and a

302 'House' group. Each group was instructed to use mental imagery strategies.

303 They were given examples of what might work (Figure 1), although the

304 participants could use their own interpretation. Specific examples for the

305 house group were 'think about your house, or a building you are familiar with

306 such as a school or church', or 'think about walking down the road looking at

307 buildings'. Specific examples for the face group were 'think of faces of people

308 you know', 'think of celebrity faces', or 'think of memorable faces you have

309 seen recently'. Both groups were instructed to pay close attention to the

310 fluctuating red bar, and to find the best way of pushing the bar up for as much

311 and as long as possible. Both groups were instructed to use whatever

312 strategy worked best, including their own, and to vary the strategy to ensure

313 continuous control of the fluctuating red bar.

314

315 Each group was unaware of the precise nature of their feedback signal.

316 During neurofeedback training, the fluctuating red bar was driven by brain

317 activity in which the signal from PPA was subtracted from FFA for the 'Face'

318 group, and the reverse subtraction (PPA – FFA) for the 'House' group.

319 Participants were trained to modulate a *differential* training signal. Therefore,

320 the 'Face group' learned to voluntarily increase the difference in BOLD

321 between FFA and PPA. In contrast, the 'House group' learned to voluntarily

322 increase the difference in BOLD between PPA and FFA.

323

324 For each group there was a *strategy-related ROI* (e.g. FFA for the Face group

325 and a *strategy-unrelated ROI* (e.g. PPA for the Face group, and vice versa for

326 the House group, Figure 2A).

327

328 **Day 5: Transfer Session**

329 After the final neurofeedback training session, there were two transfer

330 sessions, each comprising six blocks. Each block lasted 60 s and consisted of

331 an 'upregulate' period (40 s) followed by 'rest' (20 s). During upregulation,

332 participants were required to drive their brain activity 'up', using the mental

333 imagery strategies successfully used to drive the bar during neurofeedback

334 training, but now in the absence of a feedback signal.

335

336 **Day 5: Post-training BR**

337 All participants then performed post-training BR, with the same set-up

338 described for pre-training BR. Three different BR conditions were performed

339 (2 sessions each) pseudorandomised and counterbalanced across all

340 participants: (1) **Post-training BR**. The instruction was identical to the pre-

341 training BR; (2) **Post-training BR with 'concurrent trained upregulation'.**

342 Both groups were instructed to use their trained mental imagery strategies

343 that had worked best during the training sessions while simultaneously

344 performing BR; and (3) **Post-training BR with 'concurrent non-trained**

345 **mental imagery'**. Participants were instructed to use mental imagery related

346 to either houses if in the 'Face group', or faces if in the 'House group'. Mental

347 imagery was to be performed while concurrently performing BR.

348

349 <u>**Brain Imaging**</u>
350

351 Functional data was analysed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm).

352 To allow for T1 equilibration the first five images of each session were

353 discarded. Preprocessing involved bias correction, realignment of each EPI to

354 the mean EPI, unwarping, and co-registration of the functional data to the

355 structural image. Normalisation was not performed, as initial analyses were

356 performed in native space. Data was smoothed with a 6 mm FWHM Gaussian

357 kernel and high-pass filtered (128s cut-off) to remove low-frequency noise,

358 while at the same time preserving as many of the spontaneous fMRI

359 fluctuations as possible(Cordes et al., 2001). Session-specific grand mean

360 scaling was applied with no global normalisation.

361

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

363 *Area*

364

365 **Neurofeedback**

366 BOLD signals across the 9 training sessions (acquired on Days 2-4) were

367 modeled using a GLM, with regressors for each of the 9 sessions. Boxcar

368 functions were created for the six upregulation blocks, convolved with the

369 canonical HRF. Six regressors for movement and a global constant were

370 included. Beta values from the GLM were averaged across all the voxels in

371 the ROI masks (FFA and PPA ROIs based on the functional localiser). Mean

372 percentage signal change (PSC) was then calculated. For each participant,

373 the differential mean PSC between the two ROIs (i.e. strategy-related ROI

374 minus strategy-unrelated ROI) was calculated across sessions. From this, the

375 average mean PSC across participants over the training was calculated.

376

377 **Transfer Sessions**

378 Two transfer sessions were performed, with participants performing six blocks

379 of upregulation of brain activity as trained, but now in the absence of a

380 neurofeedback signal. In a similar manner to the neurofeedback sessions

381 (see above), the differential mean PSC between the two ROIs (i.e. strategy-

382 related ROI minus strategy-unrelated ROI) was calculated across sessions,

383 and from this, the average mean PSC across participants over transfer was

384 calculated.

385

386 **Binocular Rivalry**

387 Boxcar functions were created to model the onset of the BR block, convolved

388 with the canonical HRF, for each BR condition. A GLM was performed at the

389 single participant level. Beta values for each of the trained ROIs were

390 averaged for each condition and adjusted for the global brain signal. Mean

391 percentage signal change (PSC) was then calculated.

392

393 For inferential statistical analyses, SPSS 21 (IBM Corp. Armonk, USA) was

394 used to perform ANOVAs and follow-up planned paired sample t-tests, which

395 were two-tailed unless otherwise stated.

396

397 <u>**Control Experiment- Mental Imagery:**</u>
398
399
400 **Experimental outline**
401

402 Ten different participants (age range = 22-39 years, mean age 30. years, 8

403 females) were recruited for a control BR experiment. They viewed a Dell LCD

404 monitor (width: 43.5 cm; resolution: 1600 900; refresh rate: 60 Hz) from a

405 distance of 43 cm (fixed using a chin rest) through a mirror stereoscope. The

406 stereoscope reflected the left and right sides of the screen into the

407 participants' left and right eyes, so that each eye was presented with only one

408 of the two images (house or face). In order to ensure robust fusion of

409 binocular images, prior to the start of BR task, fusion was achieved for each

410 participant by slowly moving two grey squares from the edge toward the

411 centre of the screen. At the beginning of this process the participants would

412 see two squares. By the end of this process the participants would report

413 when they were seeing one square. All testing took place in a darkened room.

414 During the viewing blocks, a face stimulus and a house stimulus were

415 presented in the left and right hemi-fields respectively. The stimuli were

416 pseudorandomised with regards to which eye received the face or house

417 stimuli. Each viewing block (40 s followed by rest 20 s) was performed with a

418 new pair of stimuli from the pool of twenty stimuli. Six blocks were performed

419 per session, for three sessions. Participants were instructed to indicate a

420 perceptual shift only if the whole exemplar was perceived; any combination or

421 'patchwork' percept regardless of the predominance of the exemplar category

422 was reported as a 'mixed' percept. The participants were instructed to switch

423 as accurately and rapidly as possible between three possible button presses

424 linked to the three perceptual states (face percept, house percept, mixed

425 percept). This resulted in measures of the cumulative duration of the percept

426 throughout the BR measurement period.

427

428 BR was performed in this manner prior to and after 3 days of consecutive

429 mental imagery training (see below).

430

431 **Mental Imagery Training over 3 Days**

432

433   Participants returned to perform mental imagery training. Participants were

434   pseudorandomised into two equal groups, and were explicitly advised to use

435   mental imagery strategies that involved faces ('Face group') or house/places

436   ('House group'). Mental imagery was undertaken while viewing a LCD monitor

437   screen with a fixed horizontal black bar. They were told to imagine pushing a

438   bar above the fixed black bar, while performing their mental imagery

439   strategies. Each mental imagery training session comprised three sessions,

440   each including six blocks of 60 s with a 'perform mental imagery' period (40 s)

441   followed by 'rest' (20 s).

442

443   **Brain Structural Analysis**

444

445   The structural analysis was performed using Tensor Based Morphometry

446   (TBM), an emerging computational analysis technique(Ceccarelli et al., 2009;

447   Farbota et al., 2012; Li et al., 2009; Wang et al., 2013; Welch et al., 2013),

448   which is better suited to studies with smaller participant samples. TBM

449   enables longitudinal quantitative assessment by identifying regional structural

450   differences from the gradients of the deformation fields that nonlinearly warp

451   each individual image to the template.

452

453   For each participant, high-resolution T1 structural images were reoriented

454   placing the anterior commissure at the MNI origin. Longitudinal nonlinear

455   registration(Ashburner and Ridgway, 2012) was performed to align the two

456   time-points (before and after training) to their within-subject average,

457   characterising the relative volumetric expansion or contraction (as the

458   divergence of a velocity field) of each voxel in each time-point with respect to

459   the average. The within-subject average images were then segmented to

460   produce grey and white matter segmentations for each participant(Ashburner

461   and Friston, 2005). These segmentations were nonlinearly aligned to their

462   group-wise average using Dartel (Ashburner, 2007), and the final Dartel

463   average template was affinely registered to MNI space. The resultant

464   between-subject transformations were then used to spatially normalise the

465   divergence maps of the velocity fields, which were finally smoothed with a

466   6mm FWHM Gaussian kernel.

467

468   Divergence measures for each participant were then extracted within

469   spherical ROIs for FFA and PPA (6 mm). The spheres were centered on

470   coordinates that demonstrated the highest functional activity within the

471   localiser ROIs across training. A t-test was then performed to establish if a

472   specific brain region had changed significantly before versus after training.

473

474   **Canonical Variate Analysis**

475

476   We used a Canonical Variate Analysis (CVA) to demonstrate that measures

477   of change in brain activation and brain structure following neurofeedback

478   training predicted changes in behavioural measures. Also known as a

479   multivariate analysis of variance, or ManCova (Friston et al., 2014, 1995),

480   CVA enables statistical inferences to be made about associations between

481   the imaging data, and behavioural data that are distributed over variables. It

482 was chosen for analysis of this dataset because it can accommodate

483 statistical dependencies between multivariate predictor variables (behavioural

484 changes) and multivariate outcome variables (functional or structural

485 measures). Neither the behavioural nor imaging data had to be examined in

486 isolation, which had the advantage that distributed changes could be

487 identified, while minimising the multiple comparisons problem. The

488 behavioural changes for each participant was the change in dominance

489 duration of each the three percepts (e.g. *strategy-related* percept, *strategy-*

490 *unrelated* percept, *mixed* percept) between the pre-training BR condition and

491 post-training BR (Figure S3), and between the pre-training BR condition and

492 post-training BR with concurrent trained up-regulation (Figure S3). As the

493 behavioural and structural measures were taken prior to and immediately after

494 neurofeedback training, the functional measures for each participant were the

495 change in the different signal between the first and the last training run (e.g.

496 run 1 and run 9). The structural measures for each participant were the

497 divergence measures for each ROI, FFA and PPA (6 mm).

498

499 The objective of the CVA was to find the linear combination of outcome

500 variables that was best predicted by a linear mixture (contrast) of structural or

501 functional components. The weights of these linear combinations are called

502 canonical vectors. The canonical variates of the outcome and predictor

503 variables are the expression of each canonical vector in each subject. Other

504 quantities generated by the CVA include Bartlett's approximate chi-squared

505 statistic for Wilks' Lambda and its associated significance, or p-value, which

506 test for the significance of a linear mapping or correlation between the

507   canonical variates (in other words, if one or more pairs of canonical variates

508   show a significant statistical dependency).

509

510

511

512

513   **<u>Results</u>**
514

515   We first examined the effect of neurofeedback training on behaviour using

516   three comparisons. We compared perceptual dynamics, specifically

517   cumulative dominance durations, performed during BR before and after

518   training. We then examined the effects of learned upregulation on BR by

519   comparing pre-training BR versus post-training BR with concurrent 'learned'

520   upregulation of brain activity. For the final comparison, we examined the non-

521   trained mental imagery on BR, by comparing perceptual dynamics during pre-

522   training BR versus post-training BR with concurrent non-trained mental

523   imagery. The effects of trained mental imagery were additionally examined

524   separately – see Mental Imagery Control Experiment and Figure S1

525   (Supplementary Materials).

526

527   As the durations of the three percepts were dependent on each other, a

528   change in one percept occurred linked to changes in one or both of the other

529   percepts.

530

531   **Within Condition Comparisons:**

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

533   Comparing behavioural measures of pre-training BR and post-training BR

534   indicated an effect of training (Figure 3). Paired t-tests revealed a significant

535   reduction in the cumulative dominance durations (i.e. how long a percept type

536   was perceived) for the strategy-unrelated percept ($t(9)=2.88, p=0.02$), and a

537   significant increase in mixed percept durations ($t(9)=2.74, p=0.02$), with no

24

538   significant change in the dominance duration of the strategy-related percept

539   (t(9)=0.46,p=0.66).

540

541   *Comparison 2. Pre-training BR vs. Post-training BR with Concurrent*

542   *Trained Upregulation*

543   Paired t-tests revealed a significant reduction in the duration of the strategy-

544   unrelated percept (t(9)=4.76,p=0.001)*,* and a significant increase in the

545   duration of the mixed percept (t(9)=2.68,p=0.03). There was no significant

546   change in the dominance duration of the strategy-related percept

547   (t(9)=0.53,p=0.61) (Figure 3). The changes in BR dynamics were similar to

548   those observed with pre-training BR vs. post-training BR (Comparison 1).

549

550   *Comparison 3. Pre-training BR vs. Post-training BR with concurrent*

551   *Non-trained Mental Imagery*

552   Paired t-tests indicated a significant reduction in the duration of the strategy-

553   related percept (t(9)=2.41, p=0.04), and a significant increase in the duration

554   of the mixed percept (t(9)=2.68, p=0.03). There was no significant change in

555   the dominance duration of the strategy-unrelated percept (t(9)=0.12, p=1.74).

556

557   We further examined differences between conditions.

558

559   **Between Condition Comparisons:**

560

561   *1. Post-training BR vs. Post-training BR with Concurrent Trained*

562   *Upregulation (Comparison 1 vs. Comparison 2)*

563   There was a significantly greater reduction in the dominance duration of the

564   strategy-unrelated percept (t(9)=2.40, p=0.04) in Comparison 2 as compared

565   with Comparison 1 (Figure 3B). There were no other significant differences

566   between the two comparisons (strategy-related percept: t(9)=0.95, p=0.37;

567   mixed percept: t(9)=0.90, p=0.39).

568

569   *2. Post-training BR vs. Post-training BR with Non-trained Mental Imagery*

570   *(Comparison 1 vs. Comparison 3)*

571   There was a significantly greater reduction in the dominance duration of the

572   strategy-related percept (t(9)=3.12, p=0.01) (Figure 3B). There was also a

573   significantly greater increase in the dominance duration of the mixed percept

574   (t(9)=2.62, p=0.03). There were no significant changes in the strategy-

575   unrelated percept (t(9)=0.09, p=0.93)

576

577   *3. Post-training BR with Concurrent Trained Upregulation vs. Post-*

578   *training BR with Non-trained Mental Imagery (Comparison 2 vs.*

579   *Comparison 3)*

580   There was a trend towards reduction in the dominance duration of the

581   strategy-related percept (t(9)=2.23, p=0.05) in Comparison 3 as compared

582   with Comparison 2 (Figure 3B). The other two comparisons were not

583   significant (strategy-related percept: t(9)=0.95, p=0.37; mixed percept:

584   t(9)=1.1, p=0.30).

585

586   **Functional Changes during Neurofeedback Training**

587   To test if neurofeedback training resulted in progressive learning, we

588   examined whether participants demonstrated increased control of the

589   differential feedback signal over the three training days (Figure 2B). A

590   repeated-measures ANOVA with a factor of training day (3 levels; Days 1-3)

591   demonstrated a significant effect ($F_{(2,16)}= 3.74$, $p= 0.047$). Post-hoc t-tests

592   demonstrated a significant increase in the differential signal from Day 2

593   onwards, suggesting a learning effect (Day 1: $t(9)=0.88$, $p= 0.40$; Day 2:

594    $t(9)=3.27$ $p=0.001$; Day 3: $t(8)=2.75$, $p=0.02$).

595

596   **Functional Changes during Transfer**

597   Following neurofeedback training and prior to BR, voluntary control of brain

598   activation in the absence of neurofeedback was confirmed in a 'transfer

599   session'. Differential BOLD activation (strategy-related ROI minus strategy-

600   unrelated ROI) pooled across the two transfer sessions, revealed a significant

601   effect ($t(9)=2.38$, $p= 0.04$).

602

603   **Functional Changes during Binocular Rivalry**

604   We examined task-related BOLD signals in the trained ROIs (FFA and PPA)

605   comparing pre-training BR with post-training BR. We observed significant

606   reductions in BOLD signals in both the strategy-related ROI ($t(9)= 3.43$, $p=$

607   0.007) and strategy-unrelated ROI ($t(9)= 2.26$, $p=0.04$), when comparing pre-

608   training BR with post-training BR.

609

610   Comparing pre-training BR versus post-training BR with concurrent trained

611   upregulation, there was a significant reduction in the activation level of the

27

612 strategy-unrelated ROI (t(9)= 2.48, p= 0.03). No significant change was noted

613 for the strategy-related ROI (t(9)= 1.41, p=0.19). We performed one-tailed t-

614 tests as we had an *a priori* hypothesis that following neurofeedback training,

615 participants should be able to increase the difference in BOLD activation

616 between the two trained ROIs (Figure 4).

617

618 There were no significant changes for pre-training BR versus post-training BR

619 with concurrent non-trained mental imagery (strategy related ROI: t(9)= 0.82,

620 p= 0.44; strategy unrelated ROI: t(9)= 0.83, p= 0.43).

621

622 **Mental Imagery Control Experiment**

623

624 There was no evidence of significant changes in the cumulative dominance

625 durations of any of the three percepts (strategy-related percept t(9)= 0.74, p=

626 0.48; strategy-unrelated percept*,* t(9)= 1.00, p= 0.34; mixed percept*,* t(9)=

627 2.00, p= 0.07).

628

629 ***Between Group Comparisons with 'Mental Imagery' Control Group***

630 We performed an ANOVA with a within-subjects factor of percepts (*strategy-*

631 *related percept, strategy-unrelated percept, mixed percept*)) and a between-

632 subjects factor of group (Group 1: neurofeedback, Group 2: mental imagery).

633 There was a main effect of percept (F(2,36)= 4.64, p= 0.02). There was no

634 interaction (F(2,36)= 2.65, p= 0.08) between these two factors.

635

636 We performed a second ANOVA with a within-subjects factor of percepts

637 (*strategy-related percept, strategy-unrelated percept, mixed percept*)) and a

638 between-subjects factor of group (Group 1: neurofeedback with concurrent-

639 upregulation, Group 2: mental imagery). There was a main effect of percept

640 ($F_{(2,36)}$=6.68, p=0.003), and an interaction between percept and group

641 ($F_{(2,36)}$= 5.29, p= 0.01). Follow-up two-sample t-tests looking at changes in

642 durations of the similar percepts showed a significant difference for the

643 strategy-unrelated percept ($t_{(9)}$= 2.35, p= 0.04), but not for strategy-related

644 percept ($t_{(9)}$= 1.04, p= 0.32) or the mixed percept, ($t_{(9)}$= 2.00, p= 0.08).

645

646 We performed a further ANOVA with a within-subjects factor of percepts

647 (*strategy-related percept, strategy-unrelated percept, mixed percept*)) and a

648 between-subjects factor of group (Group 1: neurofeedback with concurrent

649 non-trained mental imagery, Group 2: mental imagery). There was a main

650 effect of percept ($F_{(2,36)}$= 6.70, p= 0.003), and an interaction between

651 percept and group ($F_{(2,36)}$= 3.63, p= 0.04). Follow-up two-sample t-tests

652 looking at changes in durations of the similar percepts showed a significant

653 difference for the mixed percept ($t_{(9)}$= 2.79, p= 0.02, but not for strategy-

654 related percept ($t_{(9)}$= 1.00, p= 0.86) or the strategy-unrelated percept, ($t_{(9)}$=

655 0.29, p= 1.14).

656

657
658

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

30

684 noted between structural changes in both ROIs and the change in BR

685 dynamics produced during concurrent trained upregulation of brain activation

686 (chi-squared value = 19.64, p= 0.03). Comparison of structural measures with

687 behavioural measures during 'simple BR' before and after training was non-

688 significant (chi-square value = 13.77, p= 0.09).

689

690 Of note, the mapping weights obtained for the behavioural measures and the

691 training-related BOLD measures were independent of the mapping weights

692 obtained for the behavioural measures and the structural measures. This is

693 because these multivariate mapping values were specific to the measures

694 used in the comparisons. Finally, the interpretation of the mapping weights in

695 relation to having a positive or negative value did not indicate a positive or

696 negative change in the values (e.g. an increase or decrease in structural

697 measures). Rather they represent a positive (or negative) contribution to the

698 mapping between the multivariate predictor variable and the outcome

699 variables.

700

Discussion

701 **Discussion**

702
703

704 Participants learned to differentially regulate the amplitude of BOLD activation

705 in two higher-order visual brain regions, FFA and PPA. This was achieved in

706 realtime, through volitional control using neurofeedback training with rt-fMRI.

707 The use of a '*differential*' training signal was implemented by showing the

708 participants a 'thermometer bar' whose size represented the difference in the

709 mean BOLD signal between the two selected brain regions. By doing this, one

710 of the brain regions acted as an internal control for the other, accounting for

711 potential confounds produced by global changes in brain activation in

712 response to effects such as arousal. Furthermore, specific behavioural effects

713 linked with the direction of change of the differential training signal were

714 obtained, providing a comparison of behavioural metrics for the training

715 effect(Thibault et al., 2018). The effect on visual perception was examined

716 with an independent BR task that employed stimuli specifically engaging

717 these stimulus-selective brain regions (face stimuli for FFA, house stimuli for

718 PPA). During BR, moment-to-moment stochastic alternations between two

719 competing visual percepts are observed, while concurrent brain activity can

720 be recorded and potentially manipulated(Blake et al., 2014; Blake and

721 Logothetis, 2002).

722

723 In this study, a change in BR perceptual dynamics was observed following

724 neurofeedback training. Perception of the stimulus linked to neurofeedback

725 training was rendered more stable e.g. strategy-related percept, with a

726 reduction in the perception of the *other* stimulus e.g. strategy-unrelated

727 percept. This behavioural change occurred when comparing pre-training BR

728 with post-training BR, and additionally when participants performed post-

729 training BR while concurrently performing learned 'upregulation' of brain

730 activity. We compared pre-training 'BR' with three post-training BR conditions:

731 'post-training BR', 'post-training BR with concurrent trained upregulation', and

732 'post-training BR with concurrent non-trained mental imagery'. The first

733 comparison, examining changes during BR before and after neurofeedback

734 training, showed altered BR dynamics; specifically a reduction in the

735 cumulative dominance duration of the strategy-unrelated percept. These

736 findings are important, as they show that neurofeedback training produced a

737 behavioural effect that was: (1) counter-intuitive in that percept durations were

738 not increased in line with the verbally instructed neurofeedback training

739 strategy, which was initially expected. Rather, percept durations *not* linked to

740 the neurofeedback training strategy (e.g. strategy-unrelated percept) were

741 reduced; (2) aligned with a longstanding finding in the field, namely Levelt's

742 second proposition (discussed below); and (3) indicative of a lack of demand

743 characteristics (see also Mental Imagery Control Experiment).

744

745 There was a significant reduction in the levels of activation in both ROIs,

746 comparing pre-training BR versus post-training BR. This linked neuroimaging

747 finding was unexpected, as the prediction from existing literature(Tong et al.,

748 1998) is that BOLD activation levels in extrastriate visual areas will reflect

749 dominance durations. The expected finding might have been that activation

750 levels would be lower in the strategy-unrelated ROI. Our findings instead

751 showed that both regions were affected by neurofeedback training, as we

752   expected given that participants trained on a differential signal involving both

753   ROIs. Both ROIs demonstrated a reduction in activation, which may reflect an

754   increase in neural efficiency as a result of more precise tuning of neural

755   representations(Gimenez et al., 2014; Haler et al., 1992; Heinzel et al., 2014;

756   Vartanian et al., 2013). The exact mechanisms underlying this gain are

757   unknown, particularly in the context of neurofeedback training and thus

758   warrants further study(Poldrack, 2015).

759

760   The purpose of the second comparison ('pre-training BR' versus 'post-training

761   BR with concurrent upregulation') was to examine if there was an effect of

762   concurrent trained modulation of brain activation on BR dynamics that was

763   additive or different to the effect of neurofeedback training alone. We

764   observed a change in BR dynamics that was similar and greater to that

765   observed for pre-training BR vs. post-training BR, in that there was *more of a*

766   *reduction* in the mean dominance duration of the strategy-unrelated percept.

767   This confirmed that the effect of trained upregulation was directly aligned with

768   the effect of neurofeedback training on BR dynamics. There was a decrease

769   in the level of BOLD activation in the strategy-unrelated ROI only, with no

770   significant change in the strategy-related ROI. Interestingly, these BOLD

771   activation changes were the same as those observed during neurofeedback

772   training (a reduction in activation levels of the strategy-unrelated ROI, Figure

773   2B). This provides further evidence for a similar mechanism underlying the

774   changes in BR dynamics following training and for those observed with

775   concurrent trained upregulation. The counter-intuitive effect of training and up-

776   regulation (during BR) on the brain activations in the two ROIs (i.e. opposite to

777 an *a priori* instruction and predicted direction of activation changes) is

778 intriguing and worthy of further investigation(Abel et al., 2015; Bueichekú et

779 al., 2016).

780 The third comparison ('pre-training BR' versus 'post-training BR with non-

781 trained mental imagery') served to assess the impact of using a differential

782 training signal, which was hypothesised to have an effect on both ROIs in all

783 participants. It additionally helped reveal the role of *non-trained* mental

784 imagery in the context of prior neurofeedback training. No significant change

785 in brain activation in either ROI was observed. However, BR dynamics

786 changed in a similar manner to the other two post-neurofeedback training BR

787 conditions, with a significant reduction in the duration of the percept not linked

788 to the training strategy used during training. This reduction was significant

789 when comparing changes in perceptual dynamics across conditions. These

790 behavioural findings would therefore suggest that neurofeedback training,

791 despite the lack of a statistically significant BOLD effect, produced a more

792 general effect on the neurobiology of the two trained ROIs. The exact nature

793 of this effect may be complex, given that behavioural changes observed for

794 this condition were opposite to the direction of neurofeedback training, but

795 nonetheless sufficient to produce an effect e.g. 'House' group participants

796 specifically underwent neurofeedback training with 'House-based' mental

797 imagery strategies, and yet they generated changes in BR dynamics simply

798 by using non-trained 'face' based mental imagery strategies during the

799 performance of BR. These behavioural findings are different from Rademaker

800 and Pearson's work, in which using mental imagery training *did not* produce

801 training-related changes in BR dominance duration. Five successive days of

802   mental imagery training had no effect on BR, with no benefit being conferred

803   by expending increased effort during mental imagery generation(Rademaker

804   and Pearson, 2012). On the other hand, Rademaker and Pearson's findings

805   are in keeping with our own mental imagery control experiment, indicating the

806   relevance of neurofeedback training. We conducted a behavioural control

807   experiment in which a different group of participants performed BR before and

808   after three consecutive days of mental imagery training, which was analogous

809   to the neurofeedback training. The training was again explicitly linked to one

810   of the two stimuli used in BR (face mental imagery for a 'Face group', house

811   mental imagery for a 'House group'). However there was no targeted training

812   strategy for the brain, unlike with the neurofeedback-trained groups. No

813   significant changes in dominance durations of any of the three percepts were

814   observed.

815

816   Taken together, these results indicate that short-term intensive training over 3

817   days on a neurofeedback BOLD signal produced by two brain regions,

818   engages and alters the function and biology of *both* regions. This is

819   specifically supported by the shift in perceptual dynamics during BR following

820   neurofeedback training, and the activation changes observed in both ROIs

821   (see Results: Comparison 1). It is further supported more broadly by the

822   behavioural changes observed in all of the post-neurofeedback training BR

823   conditions, which were not observed in the mental imagery control

824   experiment. Habes et al.(Habes et al., 2016) have previously confirmed that

825   although differential regulation of category-specific visual areas can be

826   achieved after a single day of training, a linked change in BR dynamics was

827   not produced. We therefore suggest that in order for mental imagery to

828   produce a change in perception, it must be linked with neurofeedback-led

829   learning, conducted over a period of days. This may be attributable to the

830   interposition of sleep with sequential daily training. Sleep has been directly

831   linked with the offline processing necessary for the consolidation of

832   neuroprosthetic learning(Gulati et al., 2014) and associated behavioural

833   output(Gulati et al., 2017).

834

835   Mental imagery may be utilised for perceptual learning of low-level visual

836   features, and to activate stimulus-selective cortical representations(O'Craven

837   and Kanwisher, 2000; Tartaglia et al., 2009). Similarly, rt-fMRI neurofeedback

838   together with implicit operant reinforcement has been used to unconsciously

839   train patterns of activation in primary visual brain regions(Amano et al., 2016;

840   Shibata et al., 2011) to produce perceptual and associative learning of low-

841   level visual features such as colour and orientation. However, to-date neither

842   approach has successfully yielded changes in higher-order visual perception.

843   In this study, we show that coupling explicitly instructed mental imagery with

844   rt-fMRI neurofeedback training of higher-order visual brain regions produces

845   an unconscious and targeted shift in the perceptual processing of visual

846   stimuli. This result is novel and significant in providing evidence for non-

847   invasively manipulating higher-order brain function, potentially at the level of

848   directly strengthening neural representations to alter higher-order

849   perception(Fahle, 2002; Watanabe et al., 2002, 2001). From a mechanistic

850   perspective, an interesting next step might be to test if unconsciously inducing

851   specific patterns of brain activations related to category-specific stimuli will

852 produce linked shifts in perception in a similar manner to that observed in this

853 study(Watanabe et al., 2017). This would provide more direct evidence of

854 modulating neural representations.

855

856 The observed behavioural findings may constitute a neural analogue of

857 Levelt's second proposition(Levelt, 1966), as applied to stimulus perception.

858 The original proposition (see Supplementary Discussion) was based on the

859 physical properties of visual stimuli. It was recently modified to indicate that

860 '*increasing the difference in stimulus strength between the two eyes will*

861 *primarily act to increase the average perceptual dominance duration of the*

862 *stronger stimulus'*(Brascamp et al., 2015). Our work may provide evidence for

863 a neural reformulation of BR. Participants were trained on a differential signal,

864 rather than specifically training to increase the level of activation in the

865 strategy-related ROI. During training, they appeared to reduce the level of

866 activation in the strategy-unrelated ROI across the three days, while

867 maintaining a fixed level of activation in the strategy-related ROI (Figure 2B).

868 This difference in activation levels as a result of training was maintained when

869 the participants undertook the transfer sessions, an assessment of

870 upregulation in the absence of neurofeedback. The difference in ROI

871 activation levels may have therefore led to a relative difference in the

872 strengths of the neural representations linked to the visual stimulus

873 categories. In keeping with this view, we observed a reduction in the mean

874 dominance duration of the strategy-unrelated percept. This resulted in greater

875 mean dominance durations of the strategy-related percept, corresponding to

876 the ROI with the strengthened neural representation. On the basis of this, we

877 propose a possible neural analogue of the Levelt's modified second

878 proposition as follows: '*increasing the difference in neural representation*

879 *strengths between the two brain regions linked to the two monocular visual*

880 *stimuli will primarily act to increase the average perceptual dominance of the*

881 *percept linked to the stronger neural representation'.* The effect of this would

882 be to produce unconscious perceptual biasing towards the strengthened

883 percept. This mechanism for perceptual 'shaping'(Lange et al., 2018) may

884 have real-world application in conditions requiring targeted enhancement of

885 perception such as in threat detection(Miranda et al., 2015), or therapeutically

886 to reduce unwanted or aberrant percepts(Taschereau-Dumouchel et al.,

887 2018).

888

889 Several mechanisms have been put forward to explain the neural

890 underpinnings of BR. Of note, known influences on visual perception such as

891 priming and cueing have not been shown to produce changes in BR

892 dominance durations (see also Supplementary Discussion). Neurofeedback

893 with rt-fMRI provides the most direct means of testing neuronal function

894 involved in processing visual stimuli. Using a hierarchical model of BR(Dayan,

895 1998), it may be proposed that neurofeedback training of higher order brain

896 regions strengthens neuronal representations linked to the processing of

897 specific visual stimuli, leading to unconscious perceptual biasing. Preferential

898 processing of strategy-related stimuli would result in decreased dominance

899 durations of the strategy-unrelated stimuli, as was observed here. The effect

900 of neurofeedback on BR may be further considered within a Bayesian

901 framework(Lange et al., 2018). During BR, the dominant percept at any given

902   time is maintained by the highest posterior probability, at the top of the cortical

903   hierarchy. Stimulus representations at lower levels generate error signals that

904   are compared with top-down predictions. The percept is rendered more or

905   less stable in relation to bottom-up inhibition i.e. the lower the error signal, the

906   more stable the percept(Alink et al., 2010; Hohwy et al., 2008; Summerfield

907   and Koechlin, 2008). In keeping with this, BR dynamics were shifted in the

908   direction of the information represented in the trained visual brain regions.

909   Therefore, perception of the stimulus linked to training was rendered more

910   stable, with a simultaneous reduction in the stability of the perception of the

911   other stimulus, leading to a reduction in its mean dominance duration.

912   The changes in high-level visual perception following neurofeedback training

913   in this study were associated with structural changes in the trained regions

914   (see Supplementary Materials). We used a multivariate analysis technique,

915   Canonical Variate Analysis, which can accommodate multiple measures of

916   behaviour, structure, and function to help determine the overarching effect of

917   neurofeedback training. The change in BR dynamics (i.e. cumulative

918   dominance durations) was linked with measures of structural changes in FFA,

919   and PPA (Figure S3, Supplementary Materials). These preliminary findings in

920   ten participants suggest that neurofeedback training, even over a relatively

921   short period of time (3 days) can alter perception as a result of plasticity in the

922   trained brain regions(Johansen-Berg et al., 2012; Sagi et al., 2012).

923

924   In this study, we provide a direct demonstration of the rapid changes in

925   perception and neural plasticity that can be produced by neurofeedback

926   training of higher-order visual areas using rt-fMRI. Imagery-related activation

927   in higher-order visual cortex, such as the ventral visual areas, are related to

928   semantic content, and are more flexible and abstract(Orban et al., 2014) as

929   compared to early visual cortex. Therefore, the use of higher-order visual

930   areas paired with rt-fMRI neurofeedback training may provide the most potent

931   and generalizable means of enacting a change on complex perception. Neural

932   representations that give rise to prior expectations can be directly shifted in

933   the direction of neurofeedback training, even in the presence of pre-existing

934   expectations. This could lead to targeted enhancement of specific responses

935   during discrete tasks as demonstrated here using BR, or in the reduction of

936   aberrant visual perception, such as hallucinations, for therapeutic

937   effect(Lange et al., 2018).

938

939 **<u>References</u>**

940

941

942 Abel, S., Weiller, C., Huber, W., Willmes, K., Specht, K., 2015. Therapy-

943   induced brain reorganization patterns in aphasia. Brain 138, 1097–1112.

944   doi:10.1093/brain/awv022

945 Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., Muckli, L., 2010. Stimulus

946   predictability reduces responses in primary visual cortex. J. Neurosci. 30,

947   2960–6. doi:10.1523/JNEUROSCI.3730-10.2010

948 Amano, K., Shibata, K., Kawato, M., Sasaki, Y., Watanabe, T., 2016. Learning

949   to Associate Orientation with Color in Early Visual Areas by Associative

950   Decoded fMRI Neurofeedback. Curr. Biol. 26, 1861–1866.

951   doi:10.1016/j.cub.2016.05.014

952 Ashburner, J., 2007. A fast diffeomorphic image registration algorithm.

953   Neuroimage 38, 95–113. doi:10.1016/j.neuroimage.2007.07.007

954 Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26,

955   839–51. doi:10.1016/j.neuroimage.2005.02.018

956 Ashburner, J., Ridgway, G.R., 2012. Symmetric diffeomorphic modeling of

957   longitudinal structural MRI. Front. Neurosci. 6, 197.

958   doi:10.3389/fnins.2012.00197

959 Blake, R., Brascamp, J., Heeger, D.J., Brascamp, J., 2014. Can binocular

960   rivalry reveal neural correlates of consciousness ?

961 Blake, R., Logothetis, N.K., 2002. Visual competition. Nat. Rev. Neurosci. 3,

962   13–21. doi:10.1038/nrn701

963 Blefari, M.L., Sulzer, J., Hepp-Reymond, M.-C., Kollias, S., Gassert, R., 2015.

964   Improvement in precision grip force control with self-modulation of

965   primary motor cortex during motor imagery. Front. Behav. Neurosci. 9,

966   18. doi:10.3389/fnbeh.2015.00018

967 Brascamp, J.W., Klink, P.C., Levelt, W.J.M., 2015. The 'laws' of binocular

968   rivalry: 50 years of Levelt's propositions. Vision Res. 109, 20–37.

969   doi:10.1016/j.visres.2015.02.019

970 Bueichekú, E., Miró-Padilla, A., Palomar-García, M.-Á., Ventura-Campos, N.,

971   Parcet, M.-A., Barrós-Loscertales, A., Ávila, C., 2016. Reduced posterior

972   parietal cortex activation after training on a visual search task.

973   Neuroimage 135, 204–213. doi:10.1016/j.neuroimage.2016.04.059

974 Ceccarelli, A., Rocca, M.A., Pagani, E., Falini, A., Comi, G., Filippi, M., 2009.

975   Cognitive learning is associated with gray matter changes in healthy

976   human individuals: a tensor-based morphometry study. Neuroimage 48,

977   585–9. doi:10.1016/j.neuroimage.2009.07.009

978 Cordes, D., Haughton, V.M., Arfanakis, K., Carew, J.D., Turski, P.A., Moritz,

979   C.H., Quigley, M.A., Meyerand, M.E., 2001. Frequencies Contributing to

980   Functional Connectivity in the Cerebral Cortex in "Resting-state" Data.

981   AJNR Am. J. Neuroradiol. 22, 1326–1333.

982 Dayan, P., 1998. A hierarchical model of binocular rivalry. Neural Comput. 10,

983   1119–35.

984 deCharms, R.C., Christoff, K., Glover, G.H., Pauly, J.M., Whitfield, S.,

985   Gabrieli, J.D.., 2004. Learned regulation of spatially localized brain

986   activation using real-time fMRI. Neuroimage 21, 436–443.

987   doi:10.1016/j.neuroimage.2003.08.041

988 Deichmann, R., Schwarzbauer, C., Turner, R., 2004. Optimisation of the 3D

989   MDEFT sequence for anatomical brain imaging: technical implications at

990   1.5 and 3 T. Neuroimage 21, 757–67.

991   doi:10.1016/j.neuroimage.2003.09.062

992   Eickhoff, S.B., Heim, S., Zilles, K., Amunts, K., 2006. Testing anatomically

993   specified hypotheses in functional imaging using cytoarchitectonic maps,

994   NeuroImage. doi:10.1016/j.neuroimage.2006.04.204

995   Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts,

996   K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic

997   cytoarchitectonic maps and functional imaging data, NeuroImage.

998   doi:10.1016/j.neuroimage.2004.12.034

999   Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual

1000   environment. Nature 392, 598–601. doi:10.1038/33402

1001 Fahle, M., 2002. Perceptual learning: gain without pain? Nat. Neurosci. 5,

1002   923–924. doi:10.1038/nn1002-923

1003 Farbota, K.D.M., Sodhi, A., Bendlin, B.B., McLaren, D.G., Xu, G., Rowley, H.

1004   a, Johnson, S.C., 2012. Longitudinal volumetric changes following

1005   traumatic brain injury: a tensor-based morphometry study. J. Int.

1006   Neuropsychol. Soc. 18, 1006–18. doi:10.1017/S1355617712000835

1007 Friston, K.J., Frith, C.D., Frackowiak, R.S.J., Turner, R., 1995. Characterizing

1008   Dynamic Brain Responses with fMRI: A Multivariate Approach.

1009   Neuroimage 2, 166–172. doi:10.1006/nimg.1995.1019

1010 Friston, K.J., Kahan, J., Biswal, B., Razi, A., 2014. A DCM for resting state

1011   fMRI. Neuroimage 94, 396–407. doi:10.1016/j.neuroimage.2013.12.009

1012 Gimenez, P., Bugescu, N., Black, J.M., Hancock, R., Pugh, K., Nagamine, M.,

1013   Kutner, E., Mazaika, P., Hendren, R., McCandliss, B.D., Hoeft, F., 2014.

1014   Neuroimaging correlates of handwriting quality as children learn to read

1015   and write. Front. Hum. Neurosci. 8, 155. doi:10.3389/fnhum.2014.00155

1016 Goebel, R., Esposito, F., Formisano, E., 2006. Analysis of functional image

1017   analysis contest (FIAC) data with brainvoyager QX: From single-subject

1018   to cortically aligned group general linear model analysis and self-

1019   organizing group independent component analysis. Hum. Brain Mapp.

1020   27, 392–401. doi:10.1002/hbm.20249

1021 Gulati, T., Guo, L., Ramanathan, D.S., Bodepudi, A., Ganguly, K., 2017.

1022   Neural reactivations during sleep determine network credit assignment.

1023   Nat. Neurosci. 20, 1277–1284. doi:10.1038/nn.4601

1024 Gulati, T., Ramanathan, D.S., Wong, C.C., Ganguly, K., 2014. Reactivation of

1025   emergent task-related ensembles during slow-wave sleep after

1026   neuroprosthetic learning. Nat. Neurosci. 17, 1107–13.

1027   doi:10.1038/nn.3759

1028 Habes, I., Rushton, S., Johnston, S.J., Sokunbi, M.O., Barawi, K., Brosnan,

1029   M., Daly, T., Ihssen, N., Linden, D.E.J., 2016. fMRI neurofeedback of

1030   higher visual areas and perceptual biases. Neuropsychologia 85, 208–

1031   215. doi:10.1016/j.neuropsychologia.2016.03.031

1032 Haler, R.J., Siegel, B., Tang, C., Abel, L., Buchsbaum, M.S., 1992.

1033   Intelligence and Changes in Regional Cerebral Glucose Metabolic Rate

1034   Following Learning. Intelligence 16, 415–426.

1035 Heinzel, S., Lorenz, R.C., Brockhaus, W.-R., Wustenberg, T., Kathmann, N.,

1036   Heinz, A., Rapp, M.A., 2014. Working Memory Load-Dependent Brain

1037   Response Predicts Behavioral Training Gains in Older Adults. J.

1038   Neurosci. 34, 1224–1233. doi:10.1523/JNEUROSCI.2463-13.2014

1039 Hohwy, J., Roepstorff, A., Friston, K., 2008. Predictive coding explains

1040   binocular rivalry: an epistemological review. Cognition 108, 687–701.

1041   doi:10.1016/j.cognition.2008.05.010

1042 Johansen-Berg, H., Baptista, C.S., Thomas, A.G., 2012. Human structural

1043   plasticity at record speed. Neuron 73, 1058–60.

1044   doi:10.1016/j.neuron.2012.03.001

1045 Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a

1046   module in human extrastriate cortex specialized for face perception. J.

1047   Neurosci. 17, 4302–11.

1048 Koush, Y., Zvyagintsev, M., Dyck, M., Mathiak, K. a, Mathiak, K., 2012. Signal

1049   quality and Bayesian signal processing in neurofeedback based on real-

1050   time fMRI. Neuroimage 59, 478–89.

1051   doi:10.1016/j.neuroimage.2011.07.076

1052 Lange, F.P. De, Heilbron, M., Kok, P., 2018. How Do Expectations Shape

1053   Perception ? Trends Cogn. Sci. 22, 764–779.

1054   doi:10.1016/j.tics.2018.06.002

1055 Levelt, W.J.M., 1966. The alternation process in binocular rivalry. Br. J.

1056   Psychol. 57, 225–238.

1057 Li, W., He, H., Lu, J., Lv, B., Li, M., Jin, Z., 2009. <title>Detection of whole-

1058   brain abnormalities in temporal lobe epilepsy using tensor-based

1059   morphometry with DARTEL</title> 7497, 749723-749723–6.

1060   doi:10.1117/12.833128

1061 Li, X., Hartwell, K.J., Borckardt, J., Prisciandaro, J.J., Saladin, M.E., Morgan,

1062   P.S., Johnson, K.A., LeMatty, T., Brady, K.T., George, M.S., 2013.

1063   Volitional reduction of anterior cingulate cortex activity produces

1064   decreased cue craving in smoking cessation: a preliminary real-time fMRI

1065   study. Addict. Biol. 18, 739–748. doi:10.1111/j.1369-1600.2012.00449.x

1066   McCarthy, G., Puce, A., Gore, J.C., Allison, T., 1997. Face-Specific

1067   Processing in the Human Fusiform Gyrus. J. Cogn. Neurosci. 9, 605–

1068   610. doi:10.1162/jocn.1997.9.5.605

1069   Miranda, R.A., Casebeer, W.D., Hein, A.M., Judy, J.W., Krotkov, E.P., Laabs,

1070   T.L., Manzo, J.E., Pankratz, K.G., Pratt, G.A., Sanchez, J.C., Weber,

1071   D.J., Wheeler, T.L., Ling, G.S.F., 2015. DARPA-funded efforts in the

1072   development of novel brain – computer interface technologies. J.

1073   Neurosci. Methods 244, 52–67.

1074   O'Craven, K.M., Kanwisher, N., 2000. Mental imagery of faces and places

1075   activates corresponding stiimulus-specific brain regions. J. Cogn.

1076   Neurosci. 12, 1013–23.

1077   Orban, G. a, Zhu, Q., Vanduffel, W., 2014. The transition in the ventral stream

1078   from feature to real-world entity representations. Front. Psychol. 5, 695.

1079   doi:10.3389/fpsyg.2014.00695

1080   Parker, A., Alais, D., 2007. A bias for looming stimuli to predominate in

1081   binocular rivalry. Vision Res. 47, 2661–74.

1082   doi:10.1016/j.visres.2007.06.019

1083   Poldrack, R.A., 2015. Is "efficiency" a useful concept in cognitive

1084   neuroscience? Dev. Cogn. Neurosci. 11, 12–17.

1085   doi:10.1016/j.dcn.2014.06.001

1086   Rademaker, R.L., Pearson, J., 2012. Training Visual Imagery: Improvements

1087   of Metacognition, but not Imagery Strength. Front. Psychol. 3, 224.

1088   doi:10.3389/fpsyg.2012.00224

1089   Sagi, Y., Tavor, I., Hofstetter, S., Tzur-Moryosef, S., Blumenfeld-Katzir, T.,

1090   Assaf, Y., 2012. Learning in the Fast Lane: New Insights into

1091   Neuroplasticity. Neuron 73, 1195–1203.

1092   doi:10.1016/j.neuron.2012.01.025

1093   Shibata, K., Watanabe, T., Sasaki, Y., Kawato, M., 2011. Perceptual learning

1094   incepted by decoded fMRI neurofeedback without stimulus presentation.

1095   Science 334, 1413–5. doi:10.1126/science.1212003

1096   Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock,

1097   J., Weiskopf, N., Blefari, M.L., Rana, M., Oblak, E., Birbaumer, N., Sulzer,

1098   J., 2016. Closed-loop brain training: the science of neurofeedback. Nat.

1099   Rev. Neurosci. doi:10.1038/nrn.2016.164

1100   Subramanian, L., Hindle, J. V, Johnston, S., Roberts, M. V, Husain, M.,

1101   Goebel, R., Linden, D., 2011. Real-time functional magnetic resonance

1102   imaging neurofeedback for treatment of Parkinson's disease. J. Neurosci.

1103   31, 16309–17. doi:10.1523/JNEUROSCI.3498-11.2011

1104   Summerfield, C., Koechlin, E., 2008. A neural representation of prior

1105   information during perceptual inference. Neuron 59, 336–47.

1106   doi:10.1016/j.neuron.2008.05.021

1107   Tartaglia, E.M., Bamert, L., Mast, F.W., Herzog, M.H., 2009. Human

1108   perceptual learning by mental imagery. Curr. Biol. 19, 2081–5.

1109   doi:10.1016/j.cub.2009.10.060

1110   Taschereau-Dumouchel, V., Cortese, A., Chiba, T., Knotts, J.D., Kawato, M.,

1111   Lau, H., 2018. Towards an unconscious neural reinforcement intervention

1112   for common fears. Proc. Natl. Acad. Sci. 201721572.

1113   doi:10.1073/pnas.1721572115

1114   Thibault, R.T., MacPherson, A., Lifshitz, M., Roth, R.R., Raz, A., 2018.

1115   Neurofeedback with fMRI: A critical systematic review. Neuroimage 172,

1116   786–807. doi:10.1016/j.neuroimage.2017.12.071

1117 Tong, F., Meng, M., Blake, R., 2006. Neural bases of binocular rivalry. Trends

1118   Cogn. Sci. 10, 502–511. doi:10.1016/j.tics.2006.09.003

1119 Tong, F., Nakayama, K., Vaughan, J.T., Kanwisher, N., 1998. Binocular rivalry

1120   and visual awareness in human extrastriate cortex. Neuron 21, 753–9.

1121 Vartanian, O., Jobidon, M.-E., Bouak, F., Nakashima, A., Smith, I., Lam, Q.,

1122   Cheung, B., 2013. Working memory training is associated with lower

1123   prefrontal cortex activation in a divergent thinking task. Neuroscience

1124   236, 186–194. doi:10.1016/j.neuroscience.2012.12.060

1125 Wang, Y., Yuan, L., Shi, J., Greve, A., Ye, J., Toga, A.W., Reiss, A.L.,

1126   Thompson, P.M., 2013. Applying tensor-based morphometry to

1127   parametric surfaces can improve MRI-based disease diagnosis.

1128   Neuroimage 74, 209–30. doi:10.1016/j.neuroimage.2013.02.011

1129 Watanabe, T., Náñez, J.E., Koyama, S., Mukai, I., Liederman, J., Sasaki, Y.,

1130   2002. Greater plasticity in lower-level than higher-level visual motion

1131   processing in a passive perceptual learning task. Nat. Neurosci. 5, 1003–

1132   1009. doi:10.1038/nn915

1133 Watanabe, T., Náñez, J.E., Sasaki, Y., 2001. Perceptual learning without

1134   perception. Nature 413, 844–848. doi:10.1038/35101601

1135 Watanabe, T., Sasaki, Y., Shibata, K., Kawato, M., 2017. Advances in fMRI

1136   Real-Time Neurofeedback. Trends Cogn. Sci. 21, 997–1010.

1137   doi:10.1016/j.tics.2017.09.010

1138 Weiskopf, N., Mathiak, K., Bock, S.W., Scharnowski, F., Veit, R., Grodd, W.,

1139   Goebel, R., Birbaumer, N., 2004. Principles of a brain-computer interface

49

1140   (BCI) based on real-time functional magnetic resonance imaging (fMRI).

1141   IEEE Trans. Biomed. Eng. 51, 966–70. doi:10.1109/TBME.2004.827063

1142 Welch, K. a, Moorhead, T.W., McIntosh, a M., Owens, D.G.C., Johnstone,

1143   E.C., Lawrie, S.M., 2013. Tensor-based morphometry of cannabis use on

1144   brain structure in individuals at elevated genetic risk of schizophrenia.

1145   Psychol. Med. 43, 2087–96. doi:10.1017/S0033291712002668

1146

1147

1148

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.

51

1171   J.E, Y.K, G.R.R and F.S. developed the real-time acquisition and analysis

1172   tools. J.E and E.F performed the mental imagery control experiment. A.R

1173   performed the CVA analysis. J.E and G.R wrote the paper. All authors read

1174   and commented on the manuscript**.**

1175

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'.

1188 **Figure 2A. Schematic showing group ROIs (FFA and PPA statistical masks) on**

1189 **inflated canonical brains. Activation was extracted from these regions for**

1190 **production of the differential signal for neurofeedback training. The direction**

1191 **of regulation of these ROIs was specific for each group i.e. House Group, PPA**

1192 **up/ FFA down, Face Group, FFA up/ PPA down.**

1193

1194 **Figure 2B. Mean BOLD signal changes across groups, in the strategy-related**

1195 **ROI (red) and the strategy-unrelated ROI (blue), for each of the nine training**

1196 **sessions. The green line shows the difference in mean BOLD activation**

1197 **between the two brain regions and corresponds to the neurofeedback training**

1198 **signal that participants visualised in the scanner as a fluctuating bar. Error**

1199 **bars show ±1SEM.**

1200

1201

1202 **Figure 3A. Cumulative dominance durations across participants for pre-**
1203 **training BR, and the three post-training BR sessions: Post-training, Post-**
1204 **training BR with concurrent trained upregulation, and Post-training BR with**
1205 **concurrent non-trained mental imagery. Error bars show ±1SEM. The total**
1206 **duration of each BR block was 40s.**
1207

1208 **Figure 3B. Changes in cumulative dominance durations for binocular rivalry**

1209 **(BR) sessions, showing comparisons before and after neurofeedback training**

1210 **collapsed across both groups. Error bars indicate ±1SEM**

1211

1212 **A. Pre/post training BR comparison**

1213 **B. Pre/post-training BR with concurrent training upregulation**

1214 **C. Pre/post-training BR with concurrent non-trained mental imagery**

1215

1216 ***$p<0.05$. Double **$p<0.01$. Horizontal brackets indicate significant differences in**

1217 **the changes of cumulative dominance durations ($p<0.05$) ~ over a bracket**

1218 **indicates $p=0.07$.**

1219

1220

1221 **Figure 4. BOLD activation changes in the trained ROIs, during binocular rivalry**

1222 **(BR) sessions, before and after neurofeedback training. There was a significant**

1223 **reduction in activation in both the strategy-related ROI and the strategy-**

1224 **unrelated ROI following training. When BR was performed with concurrent**

1225 **trained up-regulation, there was a significant further decrease in BOLD**

1226 **activation in the strategy-unrelated ROI only. Error bars indicate ±1SEM. (***

1227 **$p < 0.05$).**

1228

1229 **Figures**

1. **Pre-training BR**

2. **Neurofeedback training**
2 groups of 5 participants:
-FFA minus PPA signal
'face' strategies
-PPA minus FFA signal
'house' strategies

3. **Post-training BR**

a)

BR

b)

BR + Trained Upregulation

c)

BR + Non-trained Mental imagery
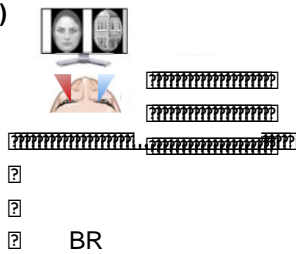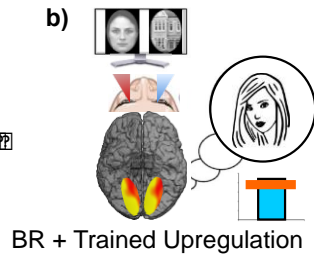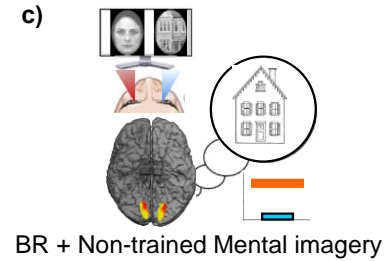
1230
1231 **Figure 1**

1232
1233
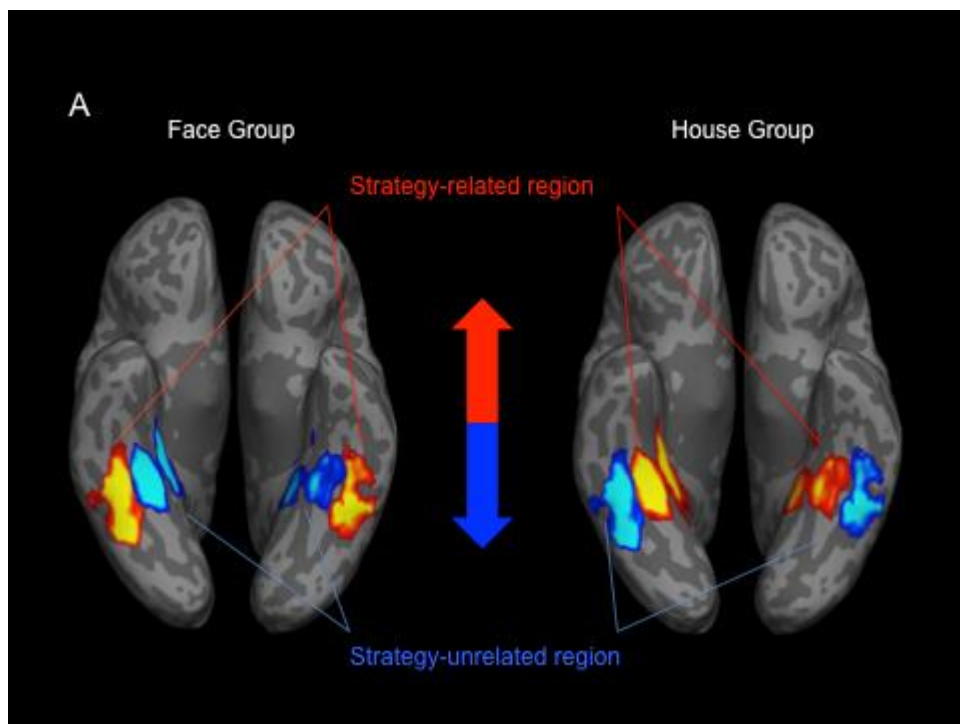
1234
1235
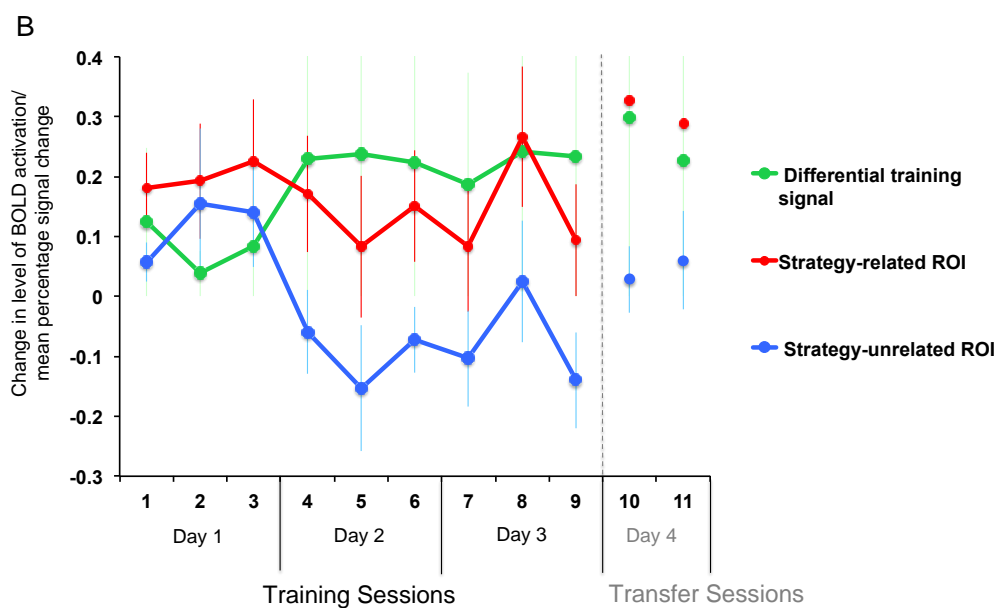
1236
1237 **Figure 2A and B.**

1238
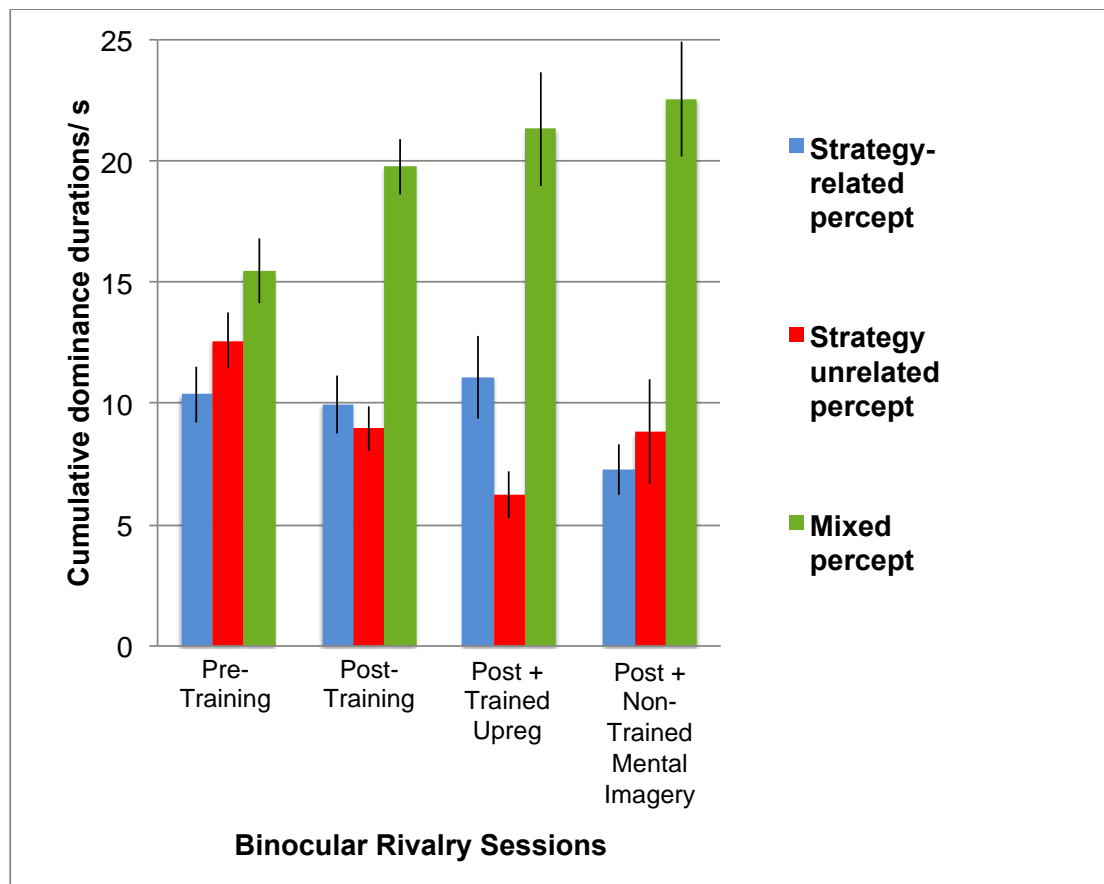
1239   **Figure 3A.**
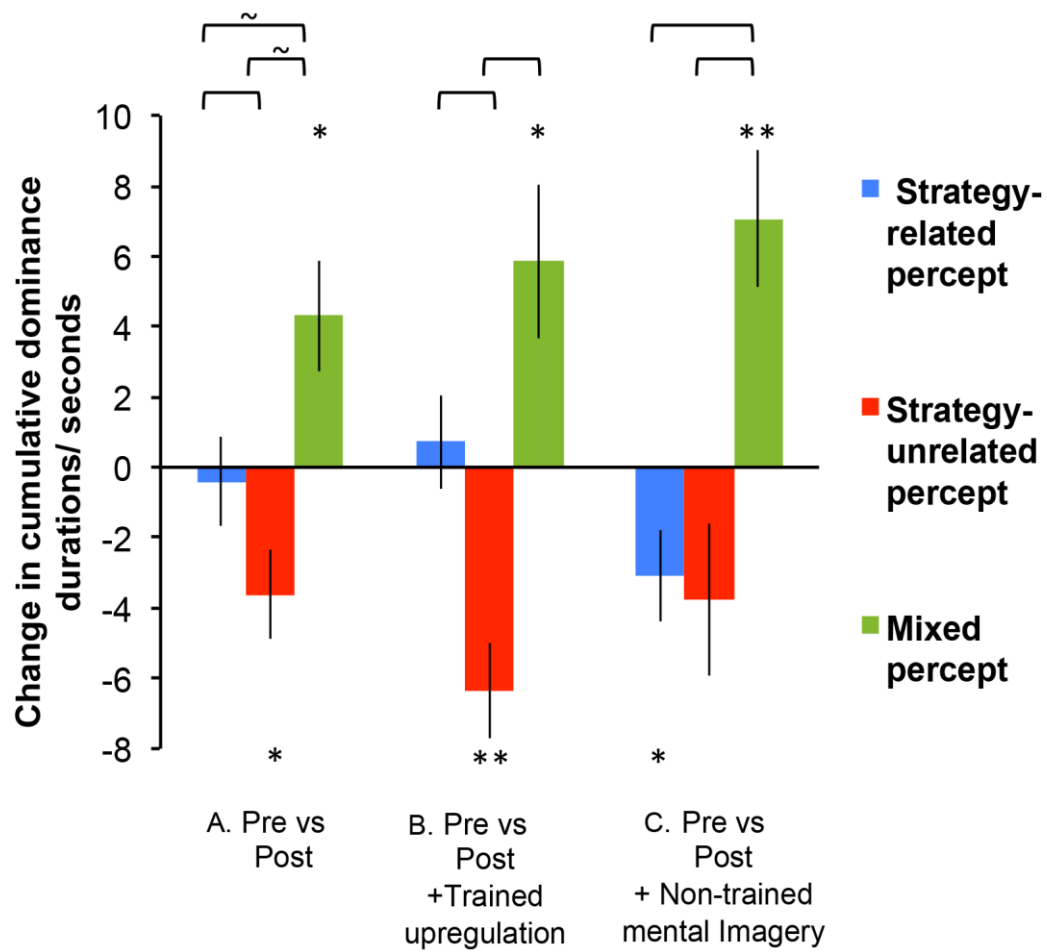
Binocular Rivalry Sessions Comparison
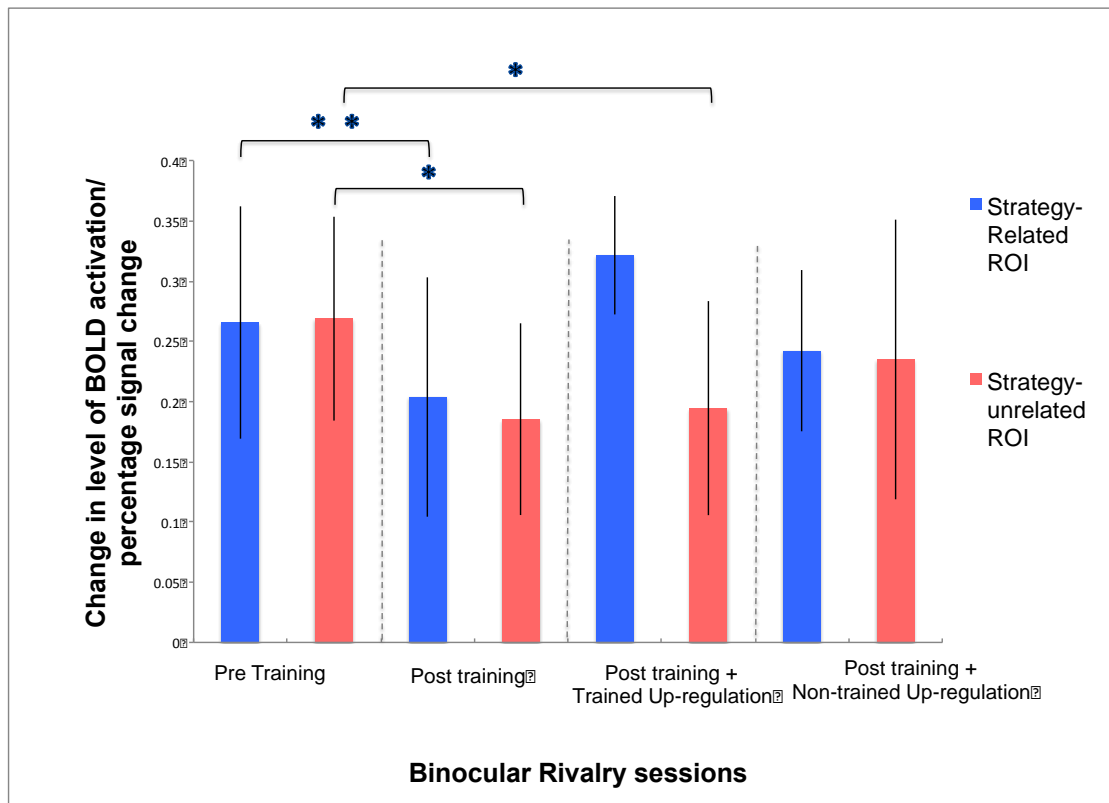
1240

1241

1242   **Figure 3B.**

1243

1244

1245
1246
1247 **Figure 4.**
1248

1255

1256

1257 **Supplementary Methods**

1258

1259 **Learning Effect across Rt-fMRI Neurofeedback Training**

1260

1261 The learning effect measures the change in BOLD activation in trained brain

1262 region/s across the neurofeedback training sessions. The mean percentage

1263 signal change (PSC) for each training run and ROI was calculated and plotted

1264 (seeFigure2B in main paper, and compare with S1, S2).

1265

1266

1267

1268

Supplementary Results

Results - Behaviour

1271
1272 **Binocular Rivalry – Durations:**

1273 See main paper.

1274

1275 **Results – Imaging**

1276

1277 *Strategy-related and Strategy-unrelated ROIs:*

1278 We first determined if the differential signal significantly changed over days

1279 across participants (see main paper, Result Section). We additionally

1280 examined the changes in the two ROIs used to produce the differential signal;

1281 the strategy-related ROI and the strategy-unrelated ROI (see Figure 2, main

1282 paper).

1283

1284 A one-way ANOVA (with 3 levels corresponding to the 3 training days)

1285 revealed a significant reduction in activation in the strategy-unrelated ROI

1286 over the 3 days of training ($F_{(2,16)}= 8.71$, $p= 0.003$). On the other hand, a

1287 one-way ANOVA for the strategy-related ROI revealed no significant change

1288 ($F_{(2,16)}= 0.33$, $p= 0.72$).

1289

1290 *Sub-groups:*

1291 To assess whether there was any difference between the face and house

1292 group during training, an ANOVA was performed on the differential training

1293 signal across the 3 training days, with a between-subjects factor with two

1294 levels (for the two sub-groups, 'Face' and 'House'). This did not reveal a

1295 significant interaction ($F_{(2,14)}=0.064$, $p=0.94$) between the two factors.

1296

1297 For neurofeedback training graphs for the two groups (mean percentage

1298 signal change over 9 sessions), please see Figures S2 and S3.

1299

1300

1301

1302 **Supplementary Discussion**

1303

1304 **Levelt's Second Proposition, 1966**

1305 Levelt's second proposition(Levelt, 1966), as applied to stimulus perception

1306 was based on the physical properties of visual stimuli and states: *"Variation of*

1307 *the stimulus strength in one eye will only influence the mean dominance*

1308 *duration of the contralateral eye and not the mean dominance duration of the*

1309 *ipsilateral eye".*

1310

1311 **Known Influences on Visual Perception**

1312 The role of 'priming' and 'cueing' might also be invoked as possible causes for

1313 the perceptual changes observed following neurofeedback training in this

1314 study. Prior presentation of a specific orientation grating can cause an

1315 increase in the perception of the identical grating during BR. However,

1316 dominance durations were unchanged(Denison et al., 2011). Similarly,

1317 exogenous cueing prior to BR can increase the probability of the predominant

1318 percept being linked to the cue. For example prior to BR, hearing sentences

1319 with the word 'face', results in FFA activation(Pelekanos et al., 2011).

1320 Nonetheless, no significant change in stimulus dominance between faces and

1321 houses on rivalry trials were observed when participants were cued with a

1322 word linked to one of the rivalrous stimuli. Dominance durations have also

1323 been demonstrated as being immune to the effects of volitional attention(Jung

1324 et al., 2016), and reflective of true differences in sensory processing(Dieter et

1325 al., 2016). It is therefore unlikely that the perceptual changes produced by

1326 neurofeedback training could be ascribed to participant expectation. Evidently,

66

1327   neither altering the level of activity in higher order brain regions involved in

1328   perception, nor applying known influences on visual perception, provide a

1329   comprehensive explanation for the lasting shifts in perceptual bistability

1330   observed following neurofeedback training in this study.

1331

1332   **Controlling the Neurofeedback Signal**

1333   With regards to the neurofeedback training signal itself (i.e. differential brain

1334   activation between two ROIs), there were five potential activation states which

1335   could increase the difference between the two brain regions (*strategy-related*

1336   *ROI minus strategy-unrelated ROI*), leading to upregulation of the training

1337   signal: These could be: (1) an increase in *strategy-related ROI*; (2) a decrease

1338   in *strategy-unrelated ROI*; (3) a combination of the two; (4) a relatively greater

1339   increase in *strategy-related ROI* as compared to *strategy-unrelated ROI*; and

1340   (5) a relatively greater decrease in the *strategy-unrelated ROI*. Based on our

1341   results (Figure 1B in main paper), the mechanism for the upregulation of the

1342   differential signal across groups during neurofeedback training appeared to be

1343   produced by maintenance of activation in the strategy-related ROI, and a

1344   reduction of activation in the strategy-unrelated ROI.

1345

1346

1347

1348

1349

Supplementary figures
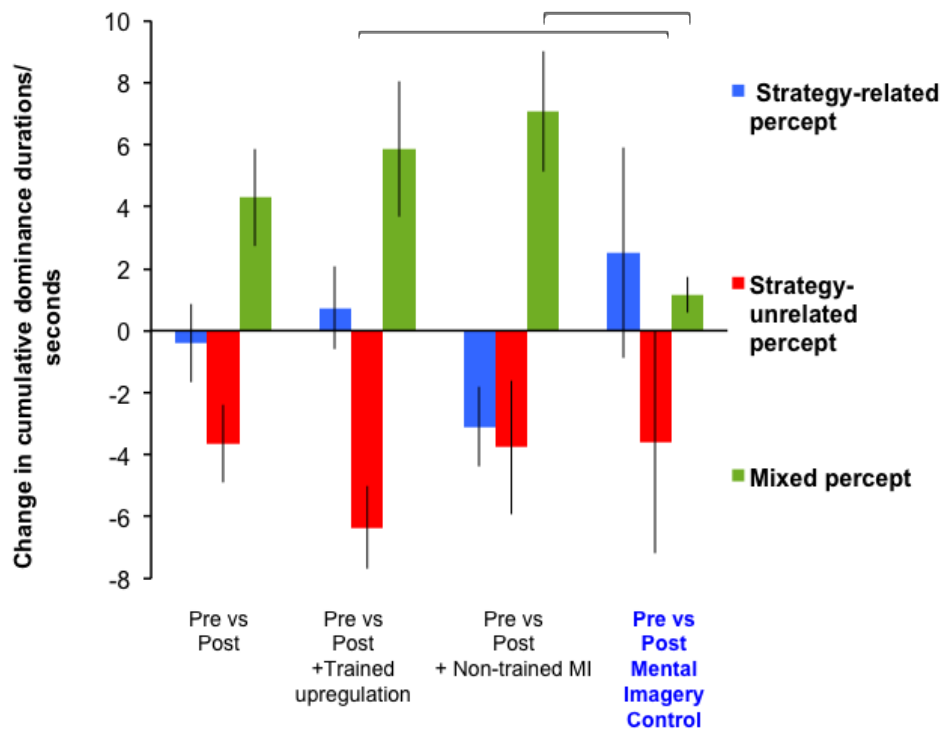
Binocular Rivalry Sessions and Groups Comparison

1351

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

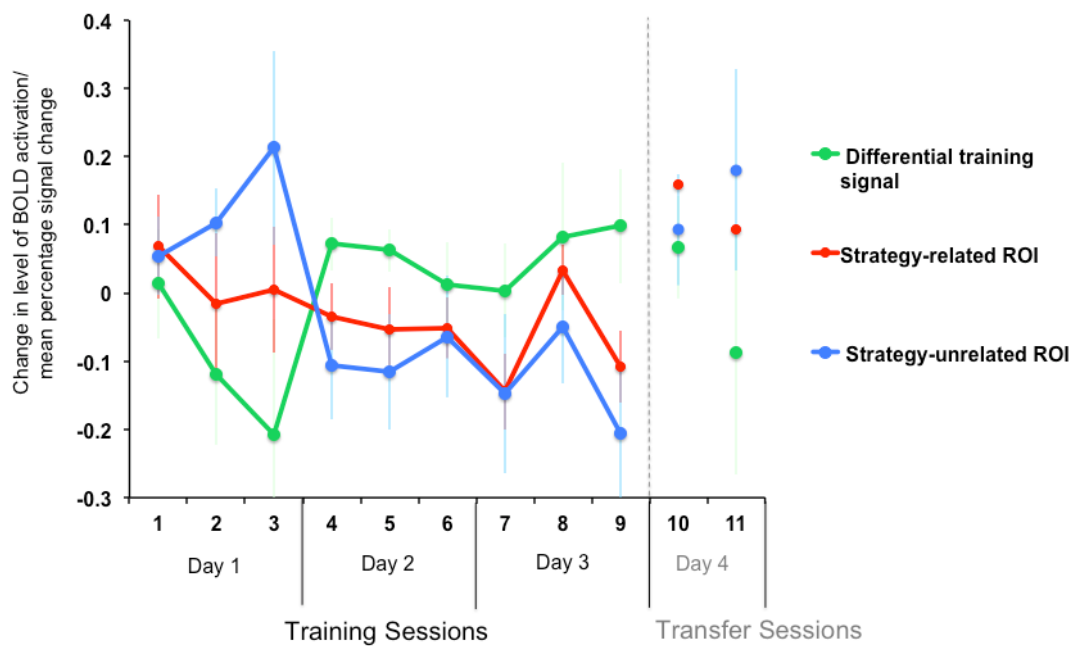**D. Pre vs. Post training BR comparison for Mental Imagery Control group**

1364

1365

1366 **Figure S2. Mean BOLD signal changes across the House group, in the strategy-**

1367 **related brain region (red) and the strategy-unrelated brain region (blue), for**

1368 **each of the nine training sessions. The green line shows the difference in mean**

1369 **BOLD activation between the two brain regions and corresponds to the**

1370 **neurofeedback training signal. Error bars show ±1SEM.**
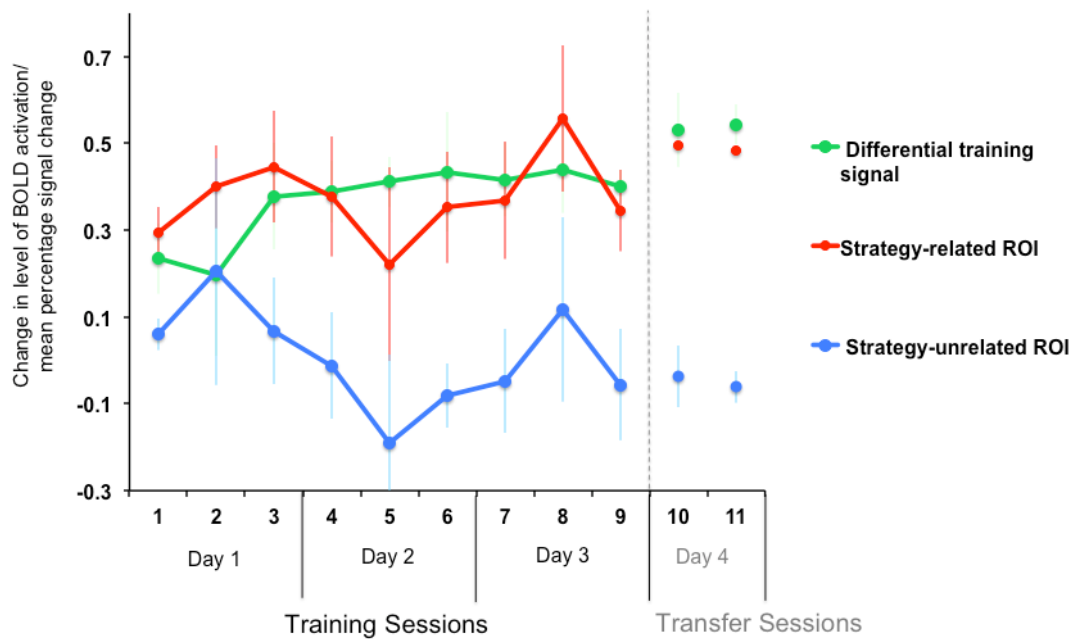
1371

1372

1373

1374

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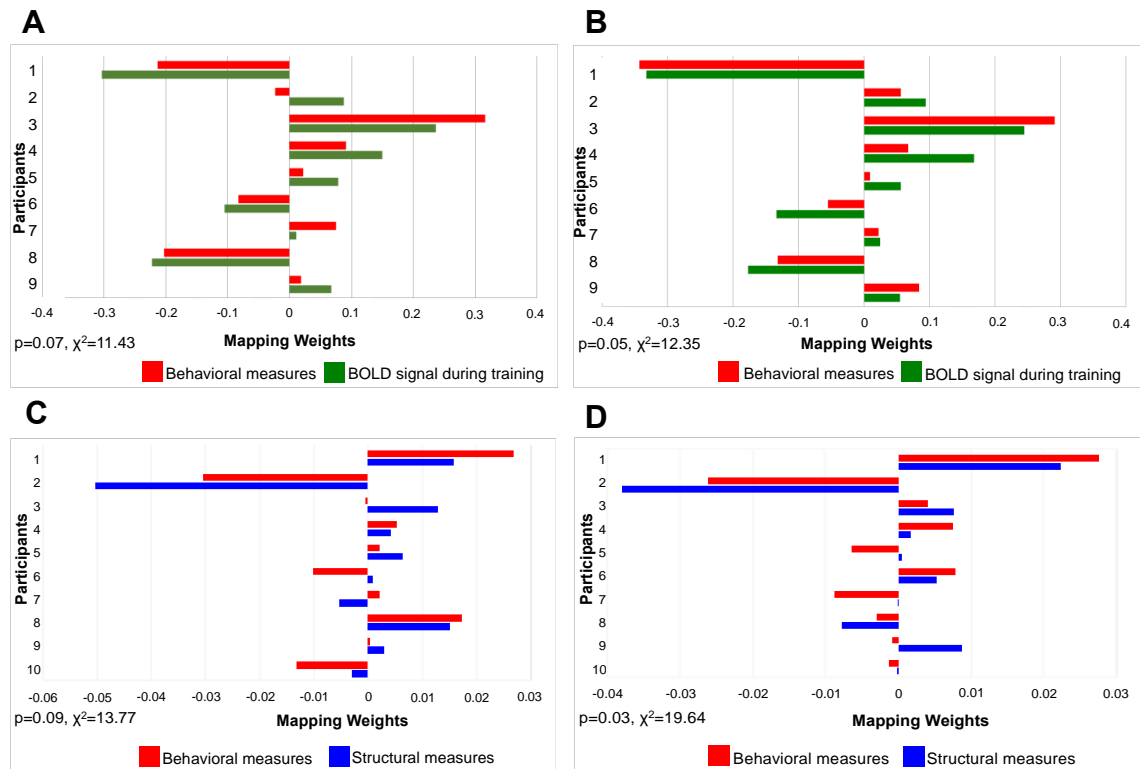
Figure S4. Canonical variate analysis illustrating the correlation between individual behaviour and physiological measures. For each participant mapping weights are shown for pairs of predictor and outcome variables. This approach aims to reveal relationships that may exist between multiple outcome variables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

1403 **individual participant BR measures (pre vs. post-training with concurrent**

1404 **trained upregulation) and functional BOLD signal changes across training.**

1405

1406 **C,D: Comparison of BR behaviour measures (i.e. durations of mixed, strategy-**

1407 **related and strategy-unrelated percepts), and structural measures from FFA**

1408 **and PPA (pre vs. post training). Participants 1-5 are 'Face Group', Participants**

1409 **6-10 are 'House Group'. C shows a non-significant relationship (p= 0.09)**

1410 **between individual participant BR measures (pre vs. post training) and**

1411 **structural measures from FFA, and PPA (pre vs. post training). D shows a**

1412 **significant relationship (p= 0.03) between individual participant BR measures**

1413 **(pre vs. post-training with concurrent trained upregulation) and structural**

1414 **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