Supplementary Information

Flexible learning of natural statistics in the human brain

D. Samuel Schwarzkopf, Jiaxiang Zhang & Zoe Kourtzi

Behavioral data for detection task during scanning: In the scanner, observers performed a target detection task. We computed hits and false alarms as well as median reaction times for hit trials. The average hit rate across conditions was high (Before training: 90.5±3.4% (SEM); After training: 94.1±2.7% (SEM)) and the false alarm rate was low (Before training: 2.4±1.0% (SEM); After training: 3.5±1.7% (SEM)). The average median reaction times on hit trials across observers were 483 ms and 477 ms before and after training, respectively. For the control experiment, observers performed the dual task with high accuracy both before (cardinal detection: 92.0±4.0% (SEM); fixation task: 86.9±5.2% (SEM)) and after training (cardinal detection: 94.9±1.6% (SEM); fixation task: 91.9±1.2% (SEM)).

Retinotopic mapping: For each individual observer we mapped the retinotopic areas using standard procedures. Eight wedge positions and eight eccentricity rings were presented for 8 s each and repeated eight times. Stimuli consisted of either gray-level natural images or black and white objects-from-texture images. Observers were instructed to fixate and performed a dimming task (i.e. detected when the fixation cross changed to the letter 'X').

Multivoxel pattern analysis (MVPA): We chose a linear Support Vector Machine (SVM) as the multivariate data classifiers. SVMs have been proposed as a powerful statistical learning method and have been widely applied in many fields (Vapnik, 1995). Sharing the common approach with other linear discrimination techniques, SVMs assign a categorical class label $y_i \in \{\pm 1\}$ to a pattern $\mathbf{x}_i$ ($i = 1, \cdots, N$, N is the number of the pattern) based on the output of the discriminant function

$$f(\mathbf{x}_i) = \mathbf{w}\mathbf{x}_i + b$$

with

$$y_i = \text{sgn}(f(\mathbf{x}_i)),$$

where the weight vector **w** and bias b define the separating hyperplane between two classes. SVMs differ from standard linear discrimination techniques in the derivation of the separating hyperplane from training patterns. SVMs maximize the margin (the distance of the nearest data point to the separating hyperplane) of separation $2/\|\mathbf{w}\|$ between two classes in feature space given that

$$y_i(\mathbf{w}\mathbf{x}_i + b) \geq 1 \text{ for all } i = 1, \cdots, N.$$

We used linear SVMs to avoid potential difficulties in the interpretation of the classification results associated with non-linear mapping from the input pattern into the feature space (Cox and Savoy, 2003; Kamitani and Tong, 2005). SVMs implement soft margin classification for noisy signals by introducing a slack variable

$$\xi_i \geq 0 \text{ for all } i = 1, \cdots, N,$$

$$y_i(\mathbf{w}\mathbf{x}_i + b) \geq 1 - \xi_i \text{ for all } i = 1, \cdots, N.$$

The separating hyperplane is obtained by minimizing the following objective function

$$E = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_i \xi_i,$$

where $C > 0$ is a penalty factor that controls the trade-off between margin maximization and training error minimization. The support vectors (SVs) are defined as the data points critical for the classification (usually near the separating hyperplane) of the training data set. Labels are assigned to independent data by comparing these data with the SVs rather than the centre of the two classes.

Supplementary Figures

Figure S1: Behavioral results: Training across sessions

Detection performance: (A) accuracy (percent correct) at zero local orientation jitter and (B) median response time plotted for each session (Pre: pre-test. Post: post-test, and each of the training sessions indicated by numbers 1-6) and stimulus condition for observers trained on orthogonal contours. Circles: collinear contours. Squares: orthogonal contours. Triangles: acute contours. Data are averaged across observers. Error bars denote ± standard error of the mean.

Figure S2: fMRI data: collinear intact vs. jittered contours

Random effects group GLM maps before (top) and after (bottom) training showing significantly stronger fMRI responses to collinear intact than jittered contours ($p<0.05$, cluster-size threshold corrected, 80 mm^2). Data is presented on a flattened reconstruction of two cortical hemispheres and retinotopic visual areas are delineated by dotted lines.

Figure S3: fMRI data: acute intact vs. jittered contours

Random effects group GLM maps before (top) and after (bottom) training showing significantly stronger fMRI responses to acute intact than jittered contours ($p<0.05$, cluster-size threshold corrected, 80 mm^2). Data is presented on a flattened reconstruction of two cortical hemispheres and retinotopic visual areas are delineated by dotted lines.

Figure S4: percent signal change from fixation baseline

Percent signal change from fixation baseline shown across ROIs for collinear and orthogonal contours before and after training. Data are averaged across observers. Error bars denote ± standard error of the mean.

Figure S5: fMRI data: orthogonal vs. acute contours.

A. Random effects group GLM maps before (top) and after (bottom) training showing significantly stronger fMRI responses to orthogonal than acute contours ($p<0.05$, cluster-size threshold corrected, 80 mm^2). Data is presented on a flattened reconstruction of two cortical hemispheres and retinotopic visual areas are delineated by dotted lines.

B. MVPA accuracy (percent correct) per ROI (voxel-pattern size of 152 voxels) for classification of orthogonal vs. acute contours. Gray bars: before training. Black bars: after training. Error bars denote standard error of the mean across observers.

Figure S6: fMRI data: collinear vs. acute contours.

A. Random effects group GLM maps before (top) and after (bottom) training showing significantly stronger fMRI responses to collinear than acute contours ($p<0.05$, cluster-size threshold corrected, 80 mm^2). Data is presented on a flattened reconstruction of two cortical hemispheres and retinotopic visual areas are delineated by dotted lines.

B. MVPA accuracy (percent correct) per ROI (voxel-pattern size of 152 voxels) for classification of collinear vs. acute contours. Gray bars: before training. Black bars: after training. Error bars denote standard error of the mean across observers.

Figure S7: Functional signal-to-noise ratio

Functional signal-to-noise ratio (fSNR) is shown before and after training across regions of interest. This is defined as the difference between the mean response to all stimuli and the response to fixation divided by the standard deviation of the mean across all stimulus conditions and fixation. No significant differences were observed between sessions (before vs. after training: $F(1,10)<1$). A significant effect of ROI ($F(4.0,40.3)=17.8$, $p<0.001$) showed the lowest fSNR in fronto-parietal regions. This cannot explain the MVPA results we observed as classification accuracy in fronto-parietal areas was high.

Figure S8: Eye movement measurements.

We recorded eye-movements for observers before (S8.1) and after training (S8.2) while they performed the experiment in the scanner. Eye movements were recorded using the ASL6000 Eye-tracker (Applied Science Laboratories, Bedford, MA). Eye tracking data was analyzed using custom Matlab (Mathworks, Natick, MA) software. For each fMRI scanning session we evaluated the mean position of fixations and the number of saccades per trial. Histograms of the horizontal (X) and vertical (Y) eye position for each stimulus were centered on fixation (zero degrees). No significant differences were observed in a one-way ANOVA across stimulus conditions in the eye position (Before training: X-position: $F(5,17)<1$, Y-position: $F(5,17)<1$; After training: X-position: $F(5,11)=1.0$, $p=0.49$, Y-position: $F(5,11)<1$) or number of saccades (Before training: $F(5,17)<1$; After training: $F(5,11)<1$). A two-way repeated measures ANOVA showed no significant differences before and after training for eye position (X-position: $F(1,2)=6.5$, $p=0.13$, Y-position: $F(1,2)=1.2$, $p=0.38$) or number of saccades ($F(1,2)<1$). These analyses suggest that it is unlikely that our results were significantly confounded by eye movements. Further, we generated an event-related eye trace for each stimulus condition and boundary that shows the change in mean eye position across the time-course of a trial. Data from each trial was brought to a common baseline (to remove drift) using the mean eye position over the 100 ms preceding stimulus onset. Plots of these traces show the event-related mean horizontal eye-position during stimulus presentation and the event-related standard deviation of horizontal eye position across trials for each stimulus condition. Mean deviations were very small and there was no evidence of systematic differences between stimulus condition or sessions.

Figure S9: individual subject data

A. Psychometric curves for two participants showing contour detection performance (percent correct) plotted as a function of local orientation jitter before training (pre-test) and after training (post-test). Error bars denote ± standard error of the mean across trials.

B, C. GLM maps before (top) and after (bottom) training showing significantly stronger fMRI responses to orthogonal intact than jittered contours ($p<0.05$, Bonferroni corrected) for each of two participants.

Supplementary Tables

Table S1: Tailarach coordinates for collinear-responsive regions

Mean Talairach coordinates (± standard deviation) across all subjects for brain regions, which showed significantly stronger activation for collinear intact than jittered contours. VOT: ventral occipito-temporal cortex; LO: lateral occipital cortex; PFs: posterior fusiform sulcus; V3B/KO: dorsal extrastriate cortex; VIPS: ventral intra-parietal sulcus; POIPS: parieto-occipital intra-parietal sulcus; DIPS: dorsal intra-parietal sulcus; PMd: pre-motor dorsal; PMv: pre-motor ventral.

	Left hemisphere						**Right hemisphere**					
	X		**Y**		**Z**		**X**		**Y**		**Z**	
VOT	**-23**	*3.6*	**-73**	*3.7*	**-13**	*6.3*	**24**	*5.2*	**-69**	*6.6*	**-13**	*3.3*
LO	**-39**	*5.0*	**-71**	*7.4*	**1**	*6.9*	**44**	*2.4*	**-60**	*4.6*	**-4**	*4.7*
PFs	**-31**	*8.7*	**-55**	*9.0*	**-12**	*2.6*	**30**	*6.4*	**-52**	*4.9*	**-12**	*3.1*
V3B/KO	**-26**	*3.1*	**-81**	*5.3*	**14**	*6.9*	**31**	*5.0*	**-80**	*2.8*	**12**	*6.3*
VIPS	**-21**	*2.9*	**-71**	*5.5*	**36**	*10.4*	**26**	*2.9*	**-70**	*4.5*	**31**	*8.8*
POIPS	**-23**	*5.8*	**-60**	*7.7*	**45**	*5.4*	**22**	*2.7*	**-62**	*5.0*	**46**	*4.3*
DIPS	**-33**	*7.0*	**-45**	*8.4*	**44**	*4.4*	**30**	*6.6*	**-46**	*4.2*	**45**	*6.5*
PMd	**-24**	*4.2*	**-7**	*5.2*	**50**	*3.2*	**27**	*3.7*	**-5**	*7.2*	**52**	*4.1*
PMv	**-46**	*3.3*	**-1**	*5.0*	**35**	*6.7*	**46**	*5.6*	**3**	*4.7*	**32**	*5.8*

Table S2: MVPA for jittered stimuli

MVPA accuracies (± standard error of the mean across observers) across ROIs (voxel pattern size of 152 voxels) for classification between the jittered versions of the different contour types (collinear vs. orthogonal, orthogonal vs. acute, collinear vs. acute).

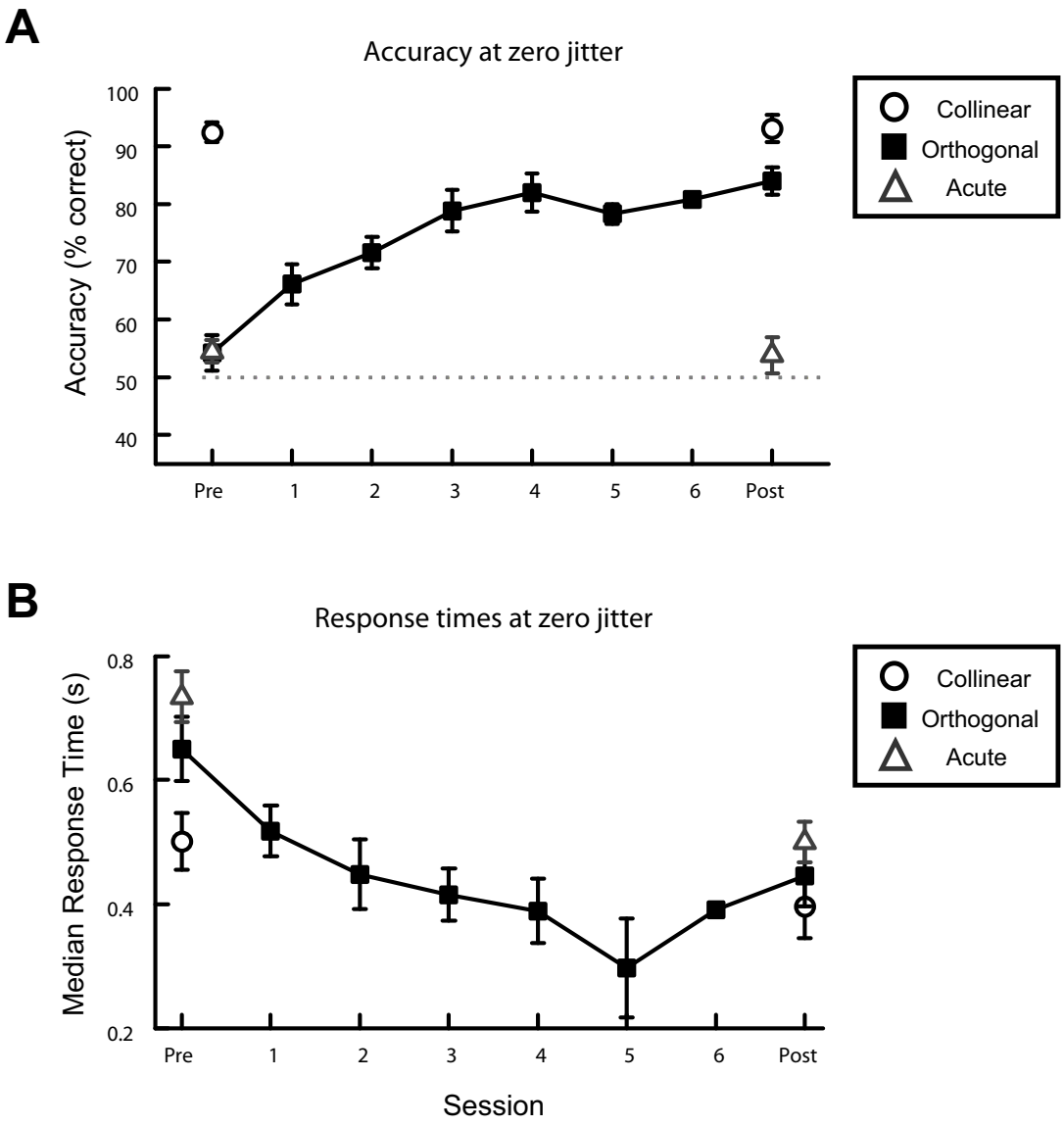
	Collinear vs Orthogonal				**Orthogonal vs Acute**				**Collinear vs Acute**			
	Before		*After*		*Before*		*After*		*Before*		*After*	
V1	**46.5**	*2.6*	**46.3**	*4.4*	**52.9**	*3.1*	**52.6**	*3.7*	**50.9**	*3.4*	**56.3**	*2.9*
V2	**50.1**	*5.0*	**50.8**	*2.7*	**50.3**	*2.7*	**51.9**	*2.9*	**54.8**	*3.9*	**53.3**	*2.6*
V3v	**42.8**	*3.6*	**42.8**	*2.9*	**48.9**	*2.5*	**50.9**	*3.5*	**51.6**	*1.9*	**49.9**	*1.1*
V4	**51.4**	*4.2*	**48.8**	*2.9*	**54.6**	*2.4*	**51.5**	*3.6*	**53.5**	*3.2*	**53.1**	*4.0*
V3d	**53.3**	*4.2*	**49.1**	*1.9*	**54.5**	*1.6*	**52.4**	*3.0*	**49.3**	*2.7*	**53.9**	*3.4*
V3A	**50.9**	*3.4*	**49.0**	*3.4*	**47.9**	*3.0*	**53.6**	*1.6*	**55.1**	*3.8*	**56.8**	*2.7*
V3B/KO	**52.5**	*3.6*	**53.9**	*2.2*	**53.4**	*4.2*	**57.8**	*2.4*	**53.3**	*2.2*	**51.1**	*3.4*
VOT	**49.4**	*3.5*	**49.0**	*2.6*	**51.9**	*2.0*	**48.1**	*2.2*	**56.6**	*3.5*	**54.8**	*1.3*
LO	**47.1**	*3.6*	**47.5**	*2.5*	**49.3**	*2.1*	**55.0**	*2.3*	**50.6**	*2.6*	**55.9**	*2.8*
PFs	**50.1**	*2.7*	**50.3**	*4.7*	**49.9**	*2.9*	**50.6**	*2.8*	**48.9**	*2.5*	**53.4**	*1.8*
VIPS	**49.1**	*2.5*	**45.3**	*3.6*	**53.0**	*5.1*	**53.7**	*3.6*	**52.8**	*4.2*	**55.4**	*1.4*
POIPS	**44.6**	*2.6*	**47.9**	*2.0*	**50.2**	*3.2*	**51.8**	*2.6*	**48.1**	*2.4*	**52.4**	*2.7*
DIPS	**48.5**	*2.2*	**46.5**	*4.8*	**47.1**	*2.3*	**49.7**	*3.5*	**47.3**	*2.8*	**46.3**	*3.6*
PMd	**51.9**	*2.7*	**47.7**	*3.8*	**48.4**	*2.8*	**46.6**	*3.8*	**47.6**	*2.3*	**52.6**	*3.5*
PMv	**47.1**	*2.5*	**51.4**	*1.7*	**47.1**	*3.2*	**53.9**	*2.4*	**43.9**	*4.2*	**55.3**	*3.0*

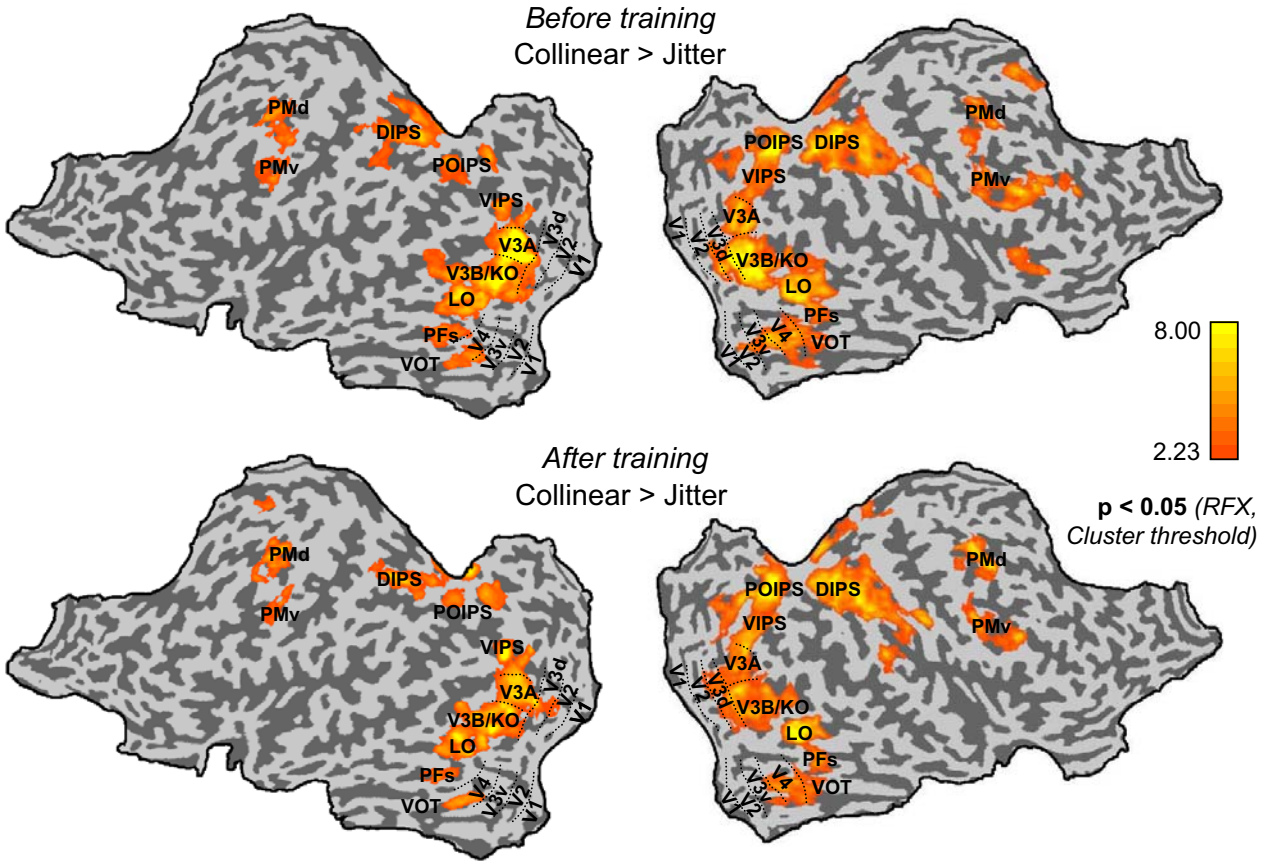
References

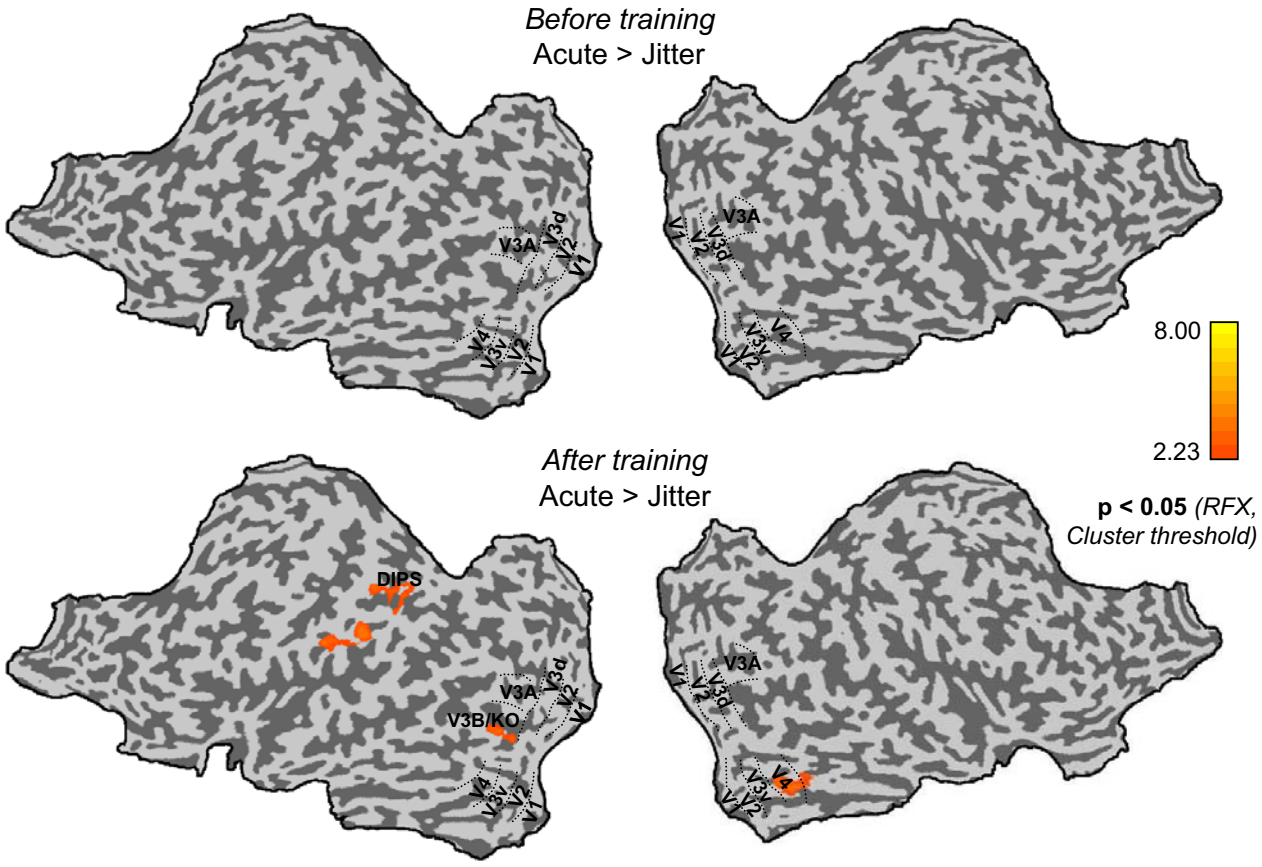
Cox DD, Savoy RL (2003) Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage 19:261-270.

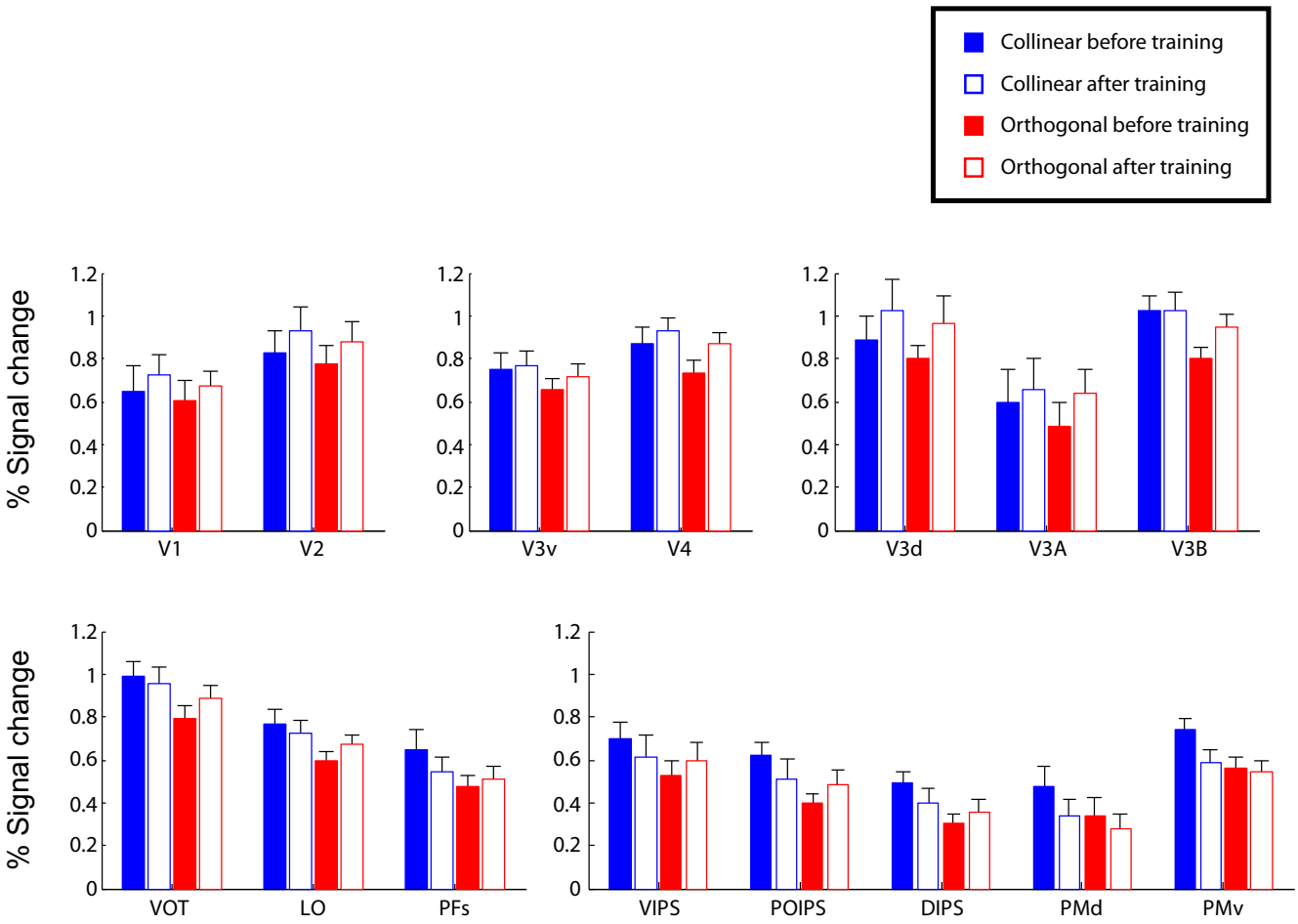
Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. Nat Neurosci 8:679-685.

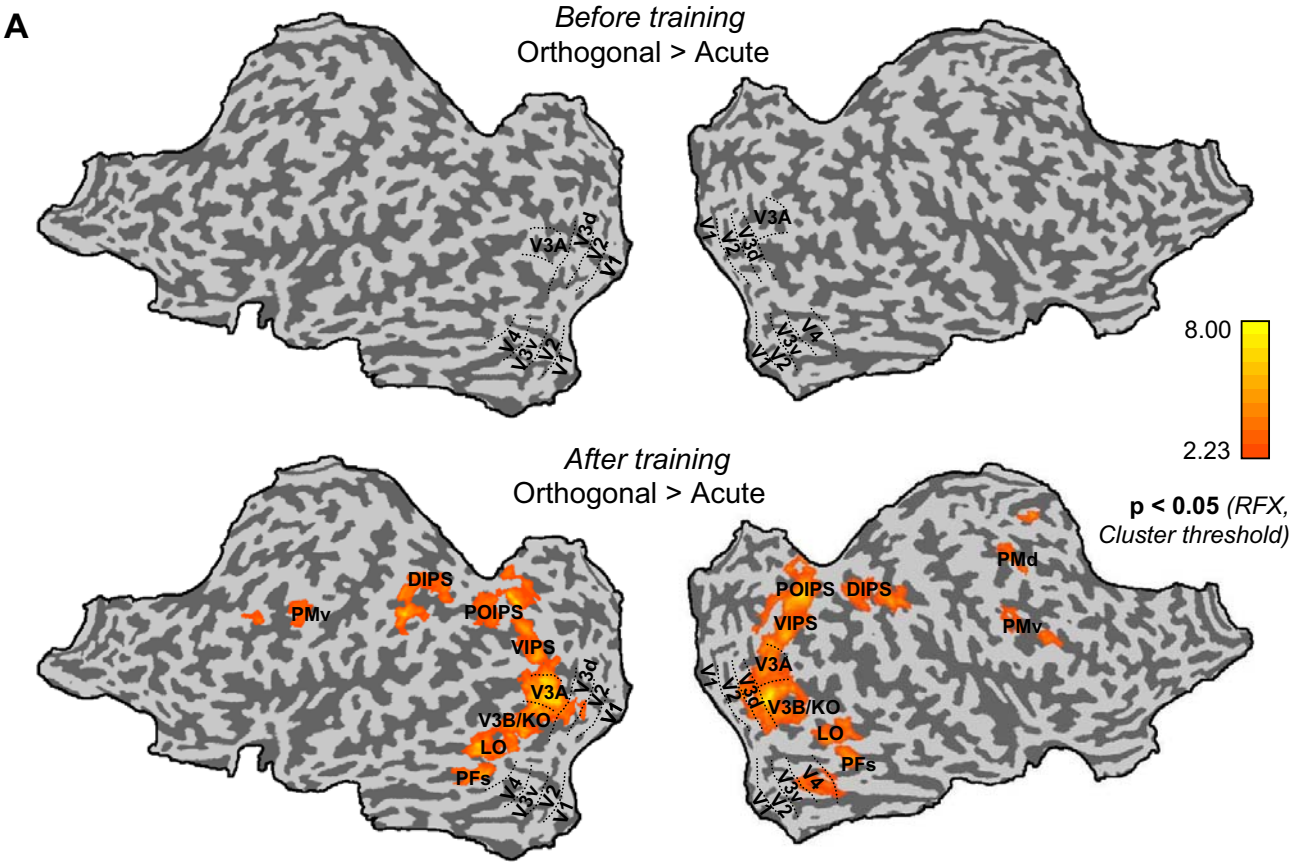
Vapnik V (1995) The Nature of Statistical Learning Theory. New York: Springer-Verlag.

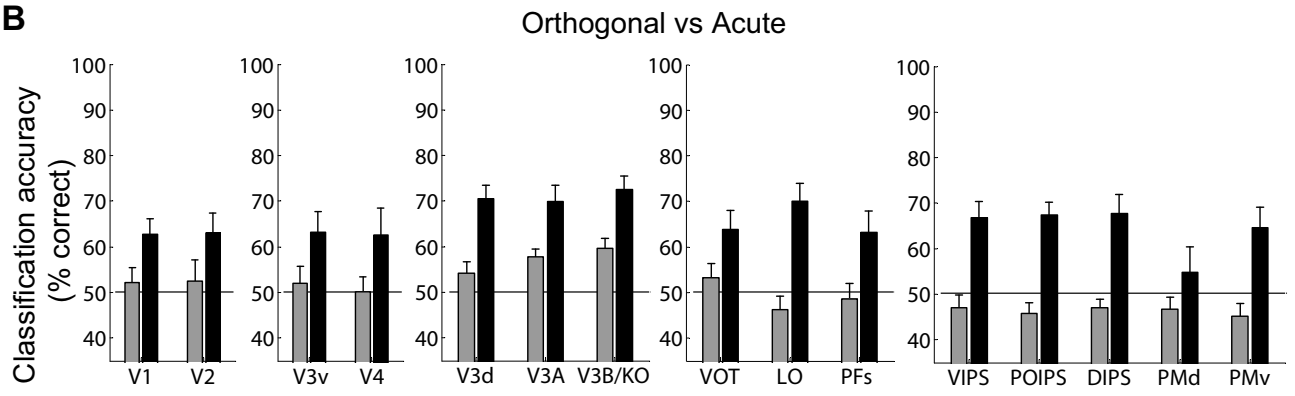

A
Accuracy at zero jitter
Accuracy (% correct)
100
90
80
70
60
50
40
Pre
1
2
3
4
5
6
Post
Collinear
Orthogonal
Acute
B
Response times at zero jitter
Median Response Time (s)
0.8
0.6
0.4
0.2
Pre
1
2
3
4
5
6
Post
Session
Collinear
Orthogonal
Acute

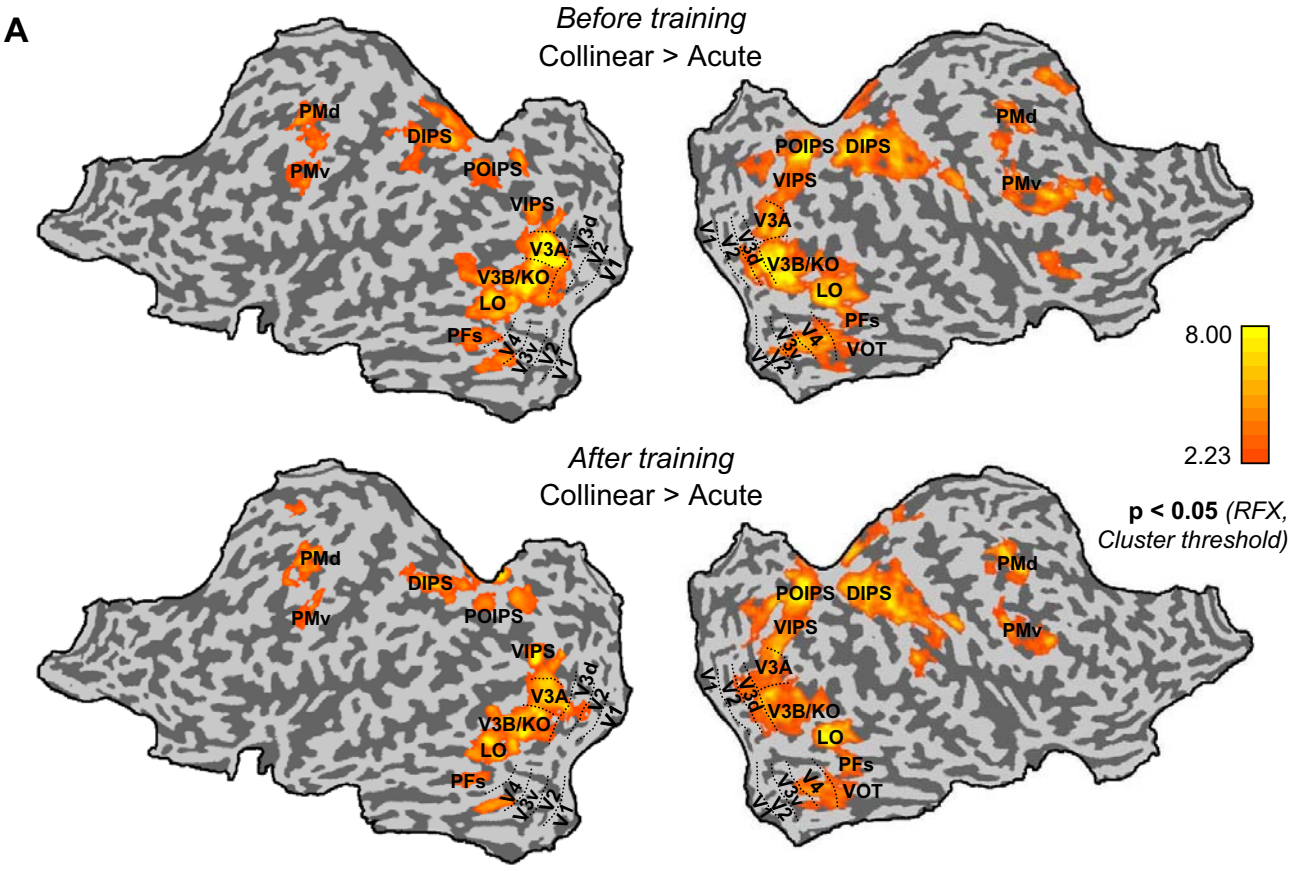

Before training
Collinear > Jitter
PMd
DIPS
PMv
POIPS
VIPS
V3A
V3B/KO
LO
PFs
VOT
V3d
V2
V1
V4
V3
8.00
2.23
After training
Collinear > Jitter
p < 0.05 (RFX, Cluster threshold)

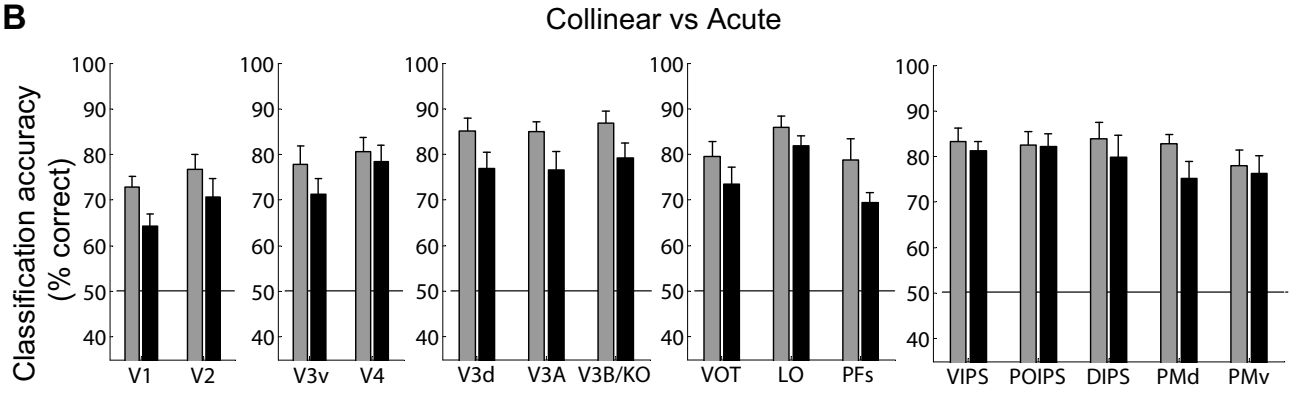

Before training
Acute > Jitter
V3A
V3d
V2
V1
V4
V3v
V2
V1
V3A
V3d
V2
V1
V4
V3v
V1/2
8.00
2.23
After training
Acute > Jitter
p < 0.05 (RFX, Cluster threshold)
DIPS
V3A
V3d
V2
V1
V3B/KO
V4
V3v
V2
V1
V3A
V3d
V2
V1
V4
V3v
V1/2

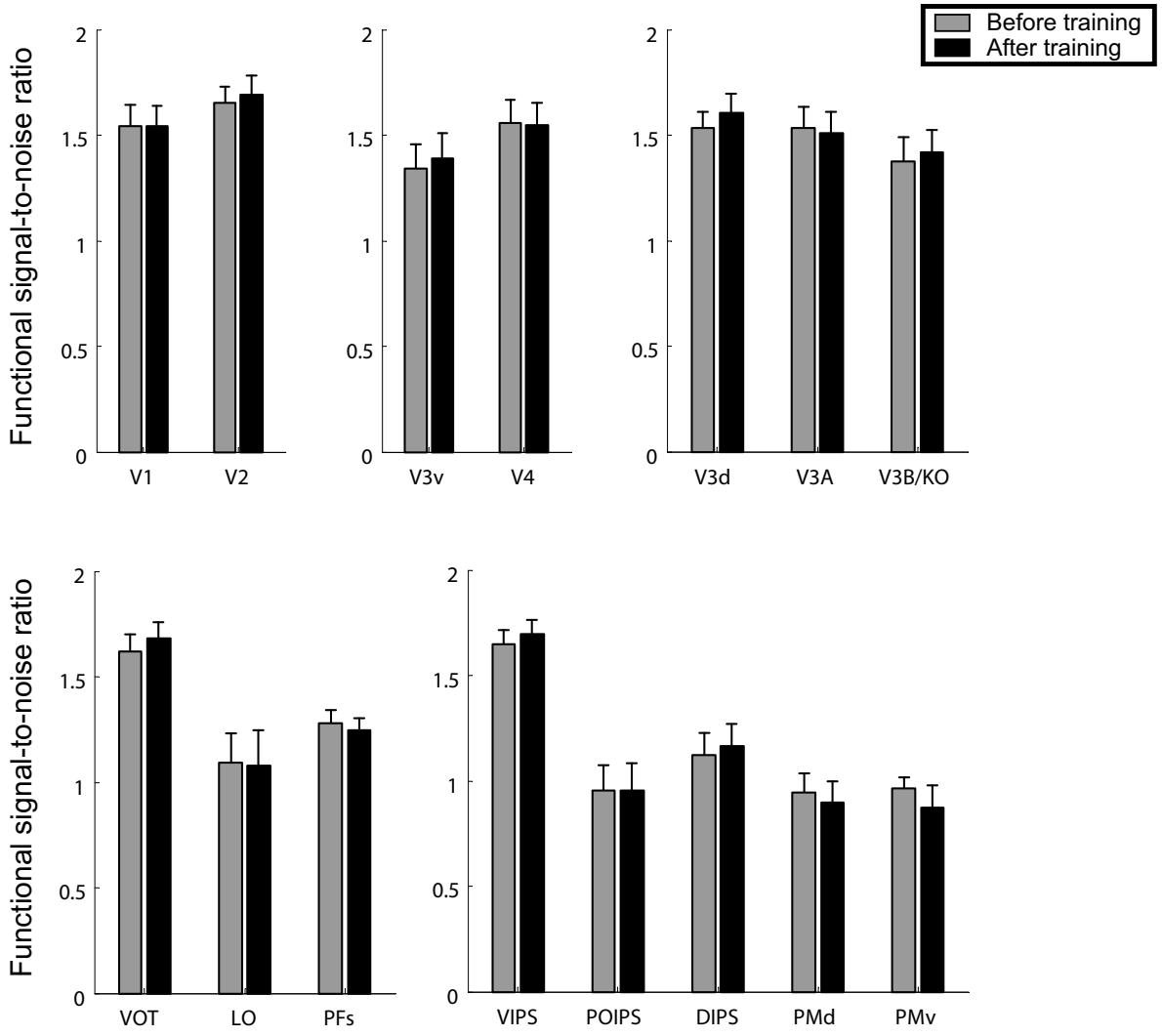

Collinear before training
Collinear after training
Orthogonal before training
Orthogonal after training
% Signal change
1.2
1
0.8
0.6
0.4
0.2
0
V1
V2
V3v
V4
V3d
V3A
V3B
% Signal change
VOT
LO
PFs
VIPS
POIPS
DIPS
PMd
PMv

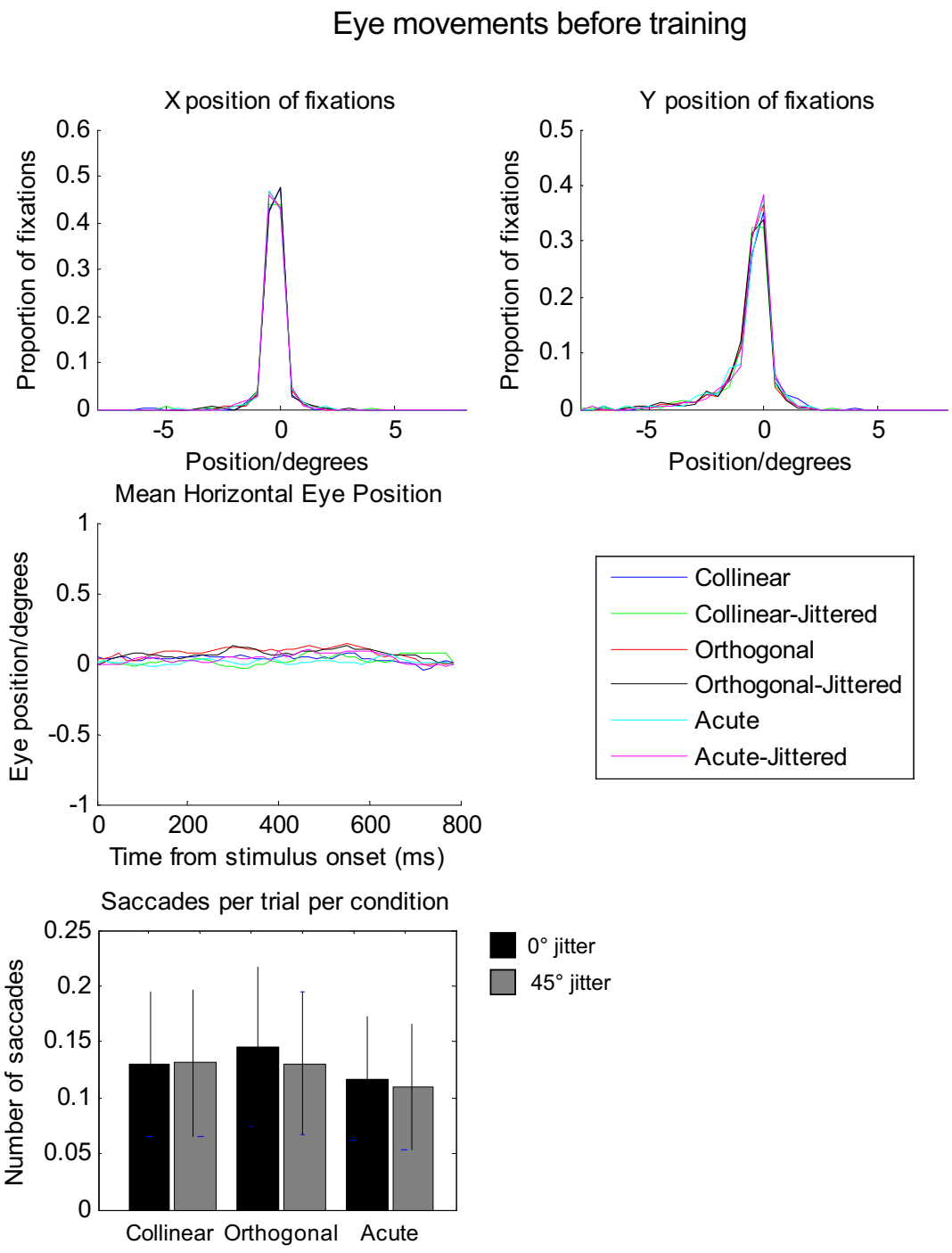

A
Before training
Orthogonal > Acute
V3A
V3d
V2
V1
V4
V3v
V2
V1
V3A
V3d
V1
V2
V4
V3v
V2
V1
8.00
2.23
After training
Orthogonal > Acute
p < 0.05 (RFX, Cluster threshold)
DIPS
PMv
POIPS
VIPS
V3A
V3d
V2
V1
V3B/KO
LO
PFs
V4
V3v
V2
V1
PMd
POIPS
DIPS
VIPS
PMv
V3A
V3d
V1
V2
V3B/KO
LO
PFs
V4
V3v
V2
V1

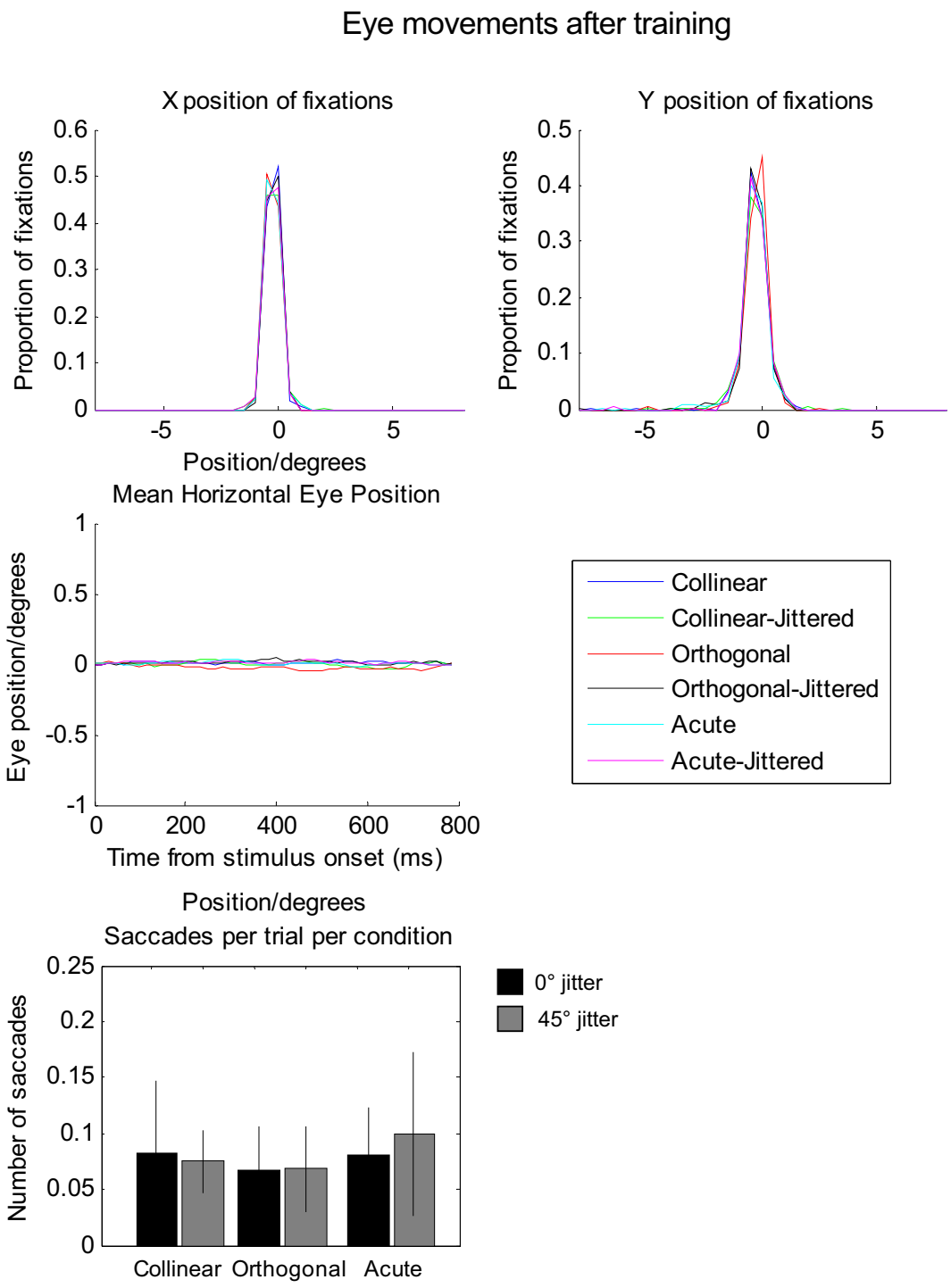

B
Orthogonal vs Acute
Classification accuracy (% correct)
100
90
80
70
60
50
40
V1
V2
V3v
V4
V3d
V3A
V3B/KO
VOT
LO
PFs
VIPS
POIPS
DIPS
PMd
PMv

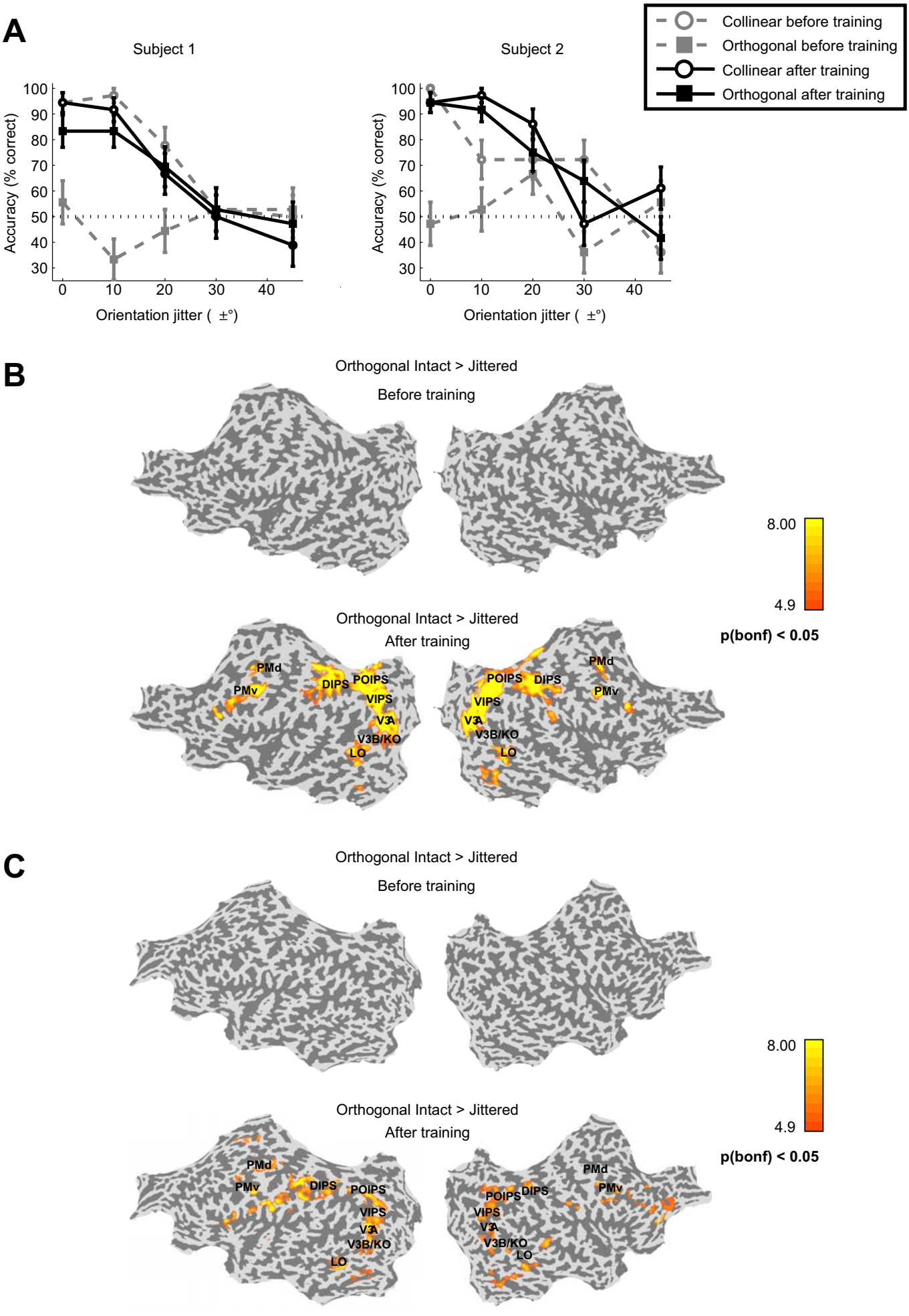

A
Before training
Collinear > Acute
PMd
DIPS
PMv
POIPS
VIPS
V3A
V3B/KO
LO
PFs
V3d
V2
V1
V4
V3v
VOT
8.00
2.23
After training
Collinear > Acute
p < 0.05 (RFX, Cluster threshold)


B
Collinear vs Acute
Classification accuracy (% correct)
100
90
80
70
60
50
40
V1
V2
V3v
V4
V3d
V3A
V3B/KO
VOT
LO
PFs
VIPS
POIPS
DIPS
PMd
PMv


Before training
After training
Functional signal-to-noise ratio
2
1.5
1
0.5
0
V1
V2
V3v
V4
V3d
V3A
V3B/KO
Functional signal-to-noise ratio
VOT
LO
PFs
VIPS
POIPS
DIPS
PMd
PMv


Eye movements before training
X position of fixations
Proportion of fixations
Position/degrees
Y position of fixations
Proportion of fixations
Position/degrees
Mean Horizontal Eye Position
Eye position/degrees
Time from stimulus onset (ms)
Collinear
Collinear-Jittered
Orthogonal
Orthogonal-Jittered
Acute
Acute-Jittered
Saccades per trial per condition
Number of saccades
0° jitter
45° jitter
Collinear
Orthogonal
Acute


Eye movements after training
X position of fixations
Proportion of fixations
0.6
0.5
0.4
0.3
0.2
0.1
0
-5
0
5
Position/degrees
Y position of fixations
Proportion of fixations
0.5
0.4
0.3
0.2
0.1
0
-5
0
5
Mean Horizontal Eye Position
Eye position/degrees
1
0.5
0
-0.5
-1
0
200
400
600
800
Time from stimulus onset (ms)
Collinear
Collinear-Jittered
Orthogonal
Orthogonal-Jittered
Acute
Acute-Jittered
Position/degrees
Saccades per trial per condition
Number of saccades
0.25
0.2
0.15
0.1
0.05
0
Collinear
Orthogonal
Acute
0° jitter
45° jitter


A
Subject 1
Subject 2
Collinear before training
Orthogonal before training
Collinear after training
Orthogonal after training
Accuracy (% correct)
Orientation jitter (±°)
B
Orthogonal Intact > Jittered
Before training
Orthogonal Intact > Jittered
After training
8.00
4.9
p(bonf) < 0.05
PMd
PMv
DIPS
POIPS
VIPS
V3A
V3B/KO
LO
C
Orthogonal Intact > Jittered
Before training
Orthogonal Intact > Jittered
After training
8.00
4.9
p(bonf) < 0.05