1 **Title:** Shape coding in occipito-temporal cortex relies on object silhouette, curvature and medial-

2 axis

3 **Authors and affiliations:** Paolo Papale[1,2,¶], Andrea Leo[1,3,¶], Giacomo Handjaras[1], Luca Cecchetti[1],

4 Pietro Pietrini[1] and Emiliano Ricciardi[1, *]

5 1. *Molecular Mind Lab, IMT School for Advanced Studies Lucca, 55100 Lucca, Italy*

6 2. *Department of Vision & Cognition, Netherlands Institute for Neuroscience (KNAW), 1105 BA*

7 *Amsterdam, Netherlands*

8 *3. Department of translational research and advanced technologies in medicine and surgery,*

9 *University of Pisa, 52100 Pisa, Italy*

10 ¶. *These authors contributed equally*

11 **Running head:** Shape coding in natural vision

12 **\*Corresponding author**: Emiliano Ricciardi: *emiliano.ricciardi@imtlucca.it*, Piazza San Francesco

13 19, 55100 Lucca, Italy

14

1

**Abstract**

15

16    Object recognition relies on different transformations of the retinal input, carried out by the

17    visual system, that range from local contrast to object shape and category. While some of those

18    transformations are thought to occur at specific stages of the visual hierarchy, the features they

19    represent are correlated (e.g., object shape and identity) and selectivity for the same feature

20    overlaps in many brain regions. This may be explained either by collinearity across

21    representations, or may instead reflect the coding of multiple dimensions by the same cortical

22    population. Moreover, orthogonal and shared components may differently impact distinctive

23    stages of the visual hierarchy. We recorded functional MRI (fMRI) activity while participants

24    passively attended to object images and employed a statistical approach that partitioned

25    orthogonal and shared object representations to reveal their relative impact on brain processing.

26    Orthogonal shape representations (silhouette, curvature and medial-axis) independently

27    explained distinct and overlapping clusters of selectivity in occitotemporal (OTC) and parietal

28    cortex. Moreover, we show that the relevance of shared representations linearly increases moving

29    from posterior to anterior regions. These results indicate that the visual cortex encodes shared

30    relations between different features in a topographic fashion and that object shape is encoded

31    along different dimensions, each representing orthogonal features.

32

**New & Noteworthy**

33

34    There are several possible ways of characterizing the shape of an object. Which shape

35    description better describes our brain responses while we passively perceive objects? Here, we

36    employed three competing shape models to explain brain representations when viewing real

37    objects. We found that object shape is encoded in a multi-dimensional fashion and thus defined by

38    the interaction of multiple features.

2

**Introduction**

Objects are often defined by their shape. However, from a theoretical perspective, the concept of shape may appear quite elusive, since the shape of an object could be defined in several possible ways. Consequently, many computational models of object shape can be constructed: closed shapes can be easily and reliably generated by combining simple elements (e.g., *geons* or medial axes), or by connecting few salient points with acute curvature, or by modulating radial frequency. Hence, we can ask which of these different descriptions is encoded in our brain and reflects the way we perceive and represent shapes. Here, we show that shape coding in the brain relies on multiple dimensions.

Indeed, all these different representations are sensitive to specific aspects of a shape and produce different ways of encoding the same object. Consider Figure 1. A silhouette descriptor (Figure 1 and 2A, red), consisting of a simple stimulus vectorization, would vary mostly depending on the global layout of a shape, but would be quite insensitive to small perturbations of the outline. Conversely, the curvature descriptor (Figure 1 and 2A, blue) would be unaffected by rotations or large transformations of a shape, but would be highly sensitive to the number of acute points on the outline. Of note, these different representations can either diverge independently to each other, as in the horizontal and vertical shapes, or can covary together, as for the shapes placed on the diagonal. Evidence from previous studies suggests that visual dimensions in natural vision are indeed highly correlated (Figure 2D; Bracci and Op de Beeck 2016; Kay 2011; Papale et al. 2019). Thus, addressing the extent to which brain regions represent different dimensions has so far proven challenging: how can we disentangle the role of different shape properties if they likely covary together?

To answer these questions, we recorded functional MRI (fMRI) activity while participants passively viewed object pictures. We employed a statistical approach that partitions orthogonal

3

63    and shared shape representations and reveals their relative impact on brain processing, while

64    controlling, at the same time, for low- and high-level features (Figure 2C; Greene et al. 2016;

65    Groen et al. 2012; Groen et al. 2018; Hebart et al. 2018; Lescroart et al. 2015; Ramakrishnan et al.

66    2014).

67         Since a recent behavioral work demonstrates that multiple dimensions are necessary to

68    model behavioral shape similarity (Morgenstern et al. 2020), we expected multiple shape

69    properties to be independently represented in the human visual cortex. At the same time, we

70    explored the way in which the brain exploits shared information between different visual

71    dimensions. In line with recent neurophysiological work (Hong et al. 2016), we expected to

72    observe a linear decrease of the role of orthogonal representations along the visual hierarchy.

73         In our approach, we tested three competing shape models. A first description was

74    computed by extracting the silhouette. The link between shape silhouette and representations in

75    the occipito-temporal cortex (OTC) has been extensively investigated in neuroimaging studies

76    (Bracci and Op de Beeck 2016; Kaiser et al. 2016; Khaligh-Razavi and Kriegeskorte 2014; Proklova

77    et al. 2016). Second, a skeletal representation of each stimulus was extracted by performing the

78    medial axis transform (Blum 1973). The spike rate of inferotemporal (IT) neurons in monkey are

79    sensitive to the medial-axis of objects (Hung et al. 2012), which also captures behavioral ratings of

80    shape similarity (Ayzenberg et al. 2019a; Ayzenberg and Lourenco 2019; Lowet et al. 2018) and

81    whose spatiotemporal association with brain activity in humans has been described in several

82    neuroimaging studies (Ayzenberg et al. 2019b; Handjaras et al. 2017; Leeds et al. 2013; Lescroart

83    and Biederman 2013; Papale et al. 2019). A third description was obtained by computing the

84    curvature distribution for each object contour. It has been showed that V4 neurons in monkey are

85    selective to a specific degree of curvature (Cadieu et al. 2007; Carlson et al. 2011; Connor et al.

86    2007). Moreover, the pivotal role of contour curvature in object perception has been extensively

4

87  demonstrated both by behavioral (Elder and Velisavljevic 2009; Lawrence et al. 2016; Long et al.

88  2017; Wolfe et al. 1992) and neuroimaging studies (Caldara et al. 2006; Long et al. 2018; Vernon et

89  al. 2016; Yue et al. 2014). In addition, we introduced two additional models: the inked area (in

90  pixels) of each stimulus as a low-level representation, and object identity, to account for high-level

91  information (Khaligh-Razavi and Kriegeskorte 2014; Kriegeskorte et al. 2008). Of note, all these

92  models differ in their complexity: the silhouette and inked area return a low-level description of

93  which parts and what extent of the visual field (and thus of the retinotopic cortex) are covered by

94  the picture of an object, respectively. Instead, medial-axis, curvature and identity provide high-

95  level, abstract representations, insensitive to viewpoint and local position.

96      As expected, we found both distinct and overlapping clusters of selectivity in OTC and in

97  parietal regions independently explained by different shape representations (i.e., silhouette,

98  curvature and medial-axis). Moreover, we showed that, while the prominence of retinotopic

99  processing on abstract information shifts abruptly moving from the occipital to the temporal

100  cortex, shared representations linearly increase from posterior to anterior regions along the visual

101  hierarchy.

102

103                              ** Figure 1 near here **

104
105                              ** Figure 2 near here **

106

107  **Materials and Methods**

108  *Subjects*

109      Seventeen subjects were enrolled for the study. Two subjects participated as pilot subjects

110  with a different version of the experimental protocol and their data were not used for the

111  subsequent analyses; data from one subject who abruptly terminated the experiment were

5

112    discarded. Fourteen subjects were further considered. The final sample comprised six females, age

113    was 24 ± 3 years, all subjects were right-handed with normal or corrected-to-normal vision and

114    were recruited among the students at the University of Pisa, Italy. Signed informed consent was

115    acquired from all subjects and all the experimental procedures were performed according to the

116    Declaration of Helsinki, under a protocol (1616/2003) approved by the Ethical Committee at the

117    University of Pisa, Italy.

118

119    *Task*

120        For this study, an event-related design was adopted. Stimuli consisted of 42 static images

121    of grayscale unfamiliar and common objects (stimuli are shown in supplementary material:

122    https://zenodo.org/record/3753126;         and       are       available       on       Zenodo:

123    https://doi.org/10.5281/zenodo.4038480), presented against a fixed gray background, with a

124    superimposed fixation cross (size: 2x2°), followed by a baseline condition characterized by a gray

125    screen with a red fixation cross.

126        A set of stimuli was selected, consisting of 24 common (animate and inanimate) and 18

127    unfamiliar objects (500x500 pixels – 20°x20°). The latter group represented existing objects that

128    combine the function and the shape of two of the common objects (e.g., a fish-shaped teapot). Of

129    note, a similar criterion has been employed for stimuli selection also in a recent study (Bracci et al.

130    2019). To build the final set of stimuli, pictures of existing objects were found on Internet, resized,

131    normalized for luminance and root-mean-square contrast.

132        Stimuli were presented with the Presentation software (Neurobehavioral Systems, Albany,

133    CA, USA) on MR-compatible goggles (VisuaStim, Resonance Technology Inc., CA, USA), with a LCD

134    at the resolution of 800x600 pixels (32°x24°). The study was organized in six runs, comprising 56

135    trials (in total: 8 repetitions for each stimulus) which consisted of 1000ms of stimulus presentation

6

136 and 7000ms of inter-stimulus interval; each run started and ended with 15 seconds of rest, to

137 estimate baseline levels of BOLD signal, and lasted 7:30 minutes. The total duration of the

138 experiment, including anatomical scans, was about 55 minutes.

139    During the functional runs, subjects were asked to fixate the cross at the center of the

140 screen. On selected trials, the cross changed its color from red to green, and subjects were asked

141 to detect such changes by pressing a key on a MR-compatible keyboard with the index finger of

142 their dominant hand. Order of trials was randomized across runs, and a different randomization

143 schema was used for each participant.

144

145 *Functional MRI data acquisition*

146    Data were acquired with a 3-Tesla GE Signa scanner (General Electric Inc., Milwaukee, WI,

147 USA) equipped with an 8-channel phased-array coil. For functional images, a gradient-echo echo-

148 planar imaging sequence (GE-EPI) was used, with TE = 40ms, TR = 2500ms, FA = 90°, 184 volumes,

149 acquisition time 7'40", including four additional dummy scans; image geometry parameters were:

150 Field-Of-View 258x258mm, 128x128 in-plane matrix, voxel size 2.03x2.03x4mm, 37 axial slices for

151 total brain coverage (z-axis extent = 148mm). To acquire detailed information of subject anatomy,

152 a 3D Fast Spoiled Gradient Echo T1-weighted sequence was also acquired (TE = 3.18ms, TR =

153 8.16ms, FA = 12°, Field-Of-View 256x256mm, 256x256 matrix size, 1mm$^3$ isotropic voxels, 256 axial

154 slices, z-axis extent 256mm).

155

156 *Functional MRI data processing*

157    Data preprocessing was carried out with AFNI (Cox 1996) and FSL 5.0 (Jenkinson et al.

158 2012). Preprocessing of functional data comprised slice timing correction with Fourier method

159 (*3dTshift*), rigid-body motion correction using the first volume of the third run as reference

7

160 (*3dvolreg*), spike removal (elimination of outliers in the functional time series, *3dDespike*),

161 smoothing with a Gaussian filter (fixed FWHM 4 mm, *3dmerge*), and scaling of BOLD time series to

162 percentage of the mean of each run (*3dTstat*, *3dcalc*). Processing of anatomical images consisted

163 of brain extraction (*bet*), segmentation for bias-field estimation and removal (*FAST*, *fslmaths*),

164 linear (*FLIRT*) and nonlinear registration (*FNIRT*) to MNI152 standard space.

165       For each subject, data from the six concatenated runs (960 time points) were used for a

166 GLM analysis (*3dDeconvolve*) with the responses for each stimulus – modeled with 1 seconds-long

167 block functions convolved with a canonical HRF – as predictors of interest, and the six motion

168 parameters plus polynomial trends up to $4^{th}$ order as predictors of no-interest.

169       Responses for individual stimuli were converted to MNI152 space by applying the

170 transformation matrices estimated as explained above, and resampled to a resolution of

171 2x2x2mm.

172

173 *Computational models*

174       Five different representations of the 42 stimuli were developed: three shape-based

175 descriptions of interest and two further models. For each model, we obtained a stimulus-specific

176 feature space, and pairwise dissimilarities between stimuli were computed to obtain a

177 representational dissimilarity matrix (RDM). Before computing shape-related information, stimuli

178 were binarized.

179       For the silhouette model, pairwise dissimilarity was computed using correlation distance (1

180 – Pearson's r). For the medial-axis model, pairwise distance between skeletal representations was

181 computed using the ShapeMatcher algorithm

182 (http://www.cs.toronto.edu/~dmac/ShapeMatcher/; (Van Eede et al. 2006)). In sum, the

183 ShapeMatcher algorithm builds the shock-graphs of each shape and then estimates their

8

184   dissimilarity as the deformation required to match different objects (Sebastian et al. 2004).

185   Curvature was computed as the chord-to-point distance (Monroy et al. 2011) in a 40-pixels

186   window. Pairwise dissimilarity was computed using correlation distance between the histograms

187   of curvature from each pair of stimuli. Finally, two further RDMs were built. For the inked-area,

188   pairwise dissimilarity was computed as the absolute difference between the number of pixels

189   covered by different objects. For identity, a binary representation was employed (Khaligh-Razavi

190   and Kriegeskorte 2014; Kriegeskorte et al. 2008). Unfamiliar stimuli were considered as belonging

191   to categories according to both their function and shape.

192

193   *Variance partitioning*

194   A variance partitioning analysis (Lescroart et al. 2015) was performed to determine

195   whether the three shape models in this study significantly explain unique components of the

196   variance of brain representations (computed using Pearson's correlation distance), as computed in

197   6 mm-radius spherical searchlights (Kriegeskorte et al. 2006). To this aim, explained variance

198   coefficient ($R^2$) was computed for each model RDM in independent linear regressions, and then all

199   the different combinations of models were tested in further multiple linear regressions. The final

200   statistic reporting the partial goodness of fit for unique and shared components was computed

201   following the work by Nimon and colleagues (2008). For example, the unique variance explained

202   by the curvature model in a specific searchlight was determined as the difference between the

203   full-model $R^2$ and the variance explained by the combination of all other models (i.e., $R^2_{curvature} =$

204   $R^2_{full} - R^2_{silhouette + medial-axis + inked\ area + identity}$). In the context of multiple linear regression, this

205   approach is better known as 'commonality analysis' (Nimon and Oswald 2013), and its popularity

206   is growing in neuroimaging (de Heer et al. 2017; Groen et al. 2018; Lescroart et al. 2015).

9

207       It should be noted that the orthogonal/unique partitions are not strictly uncorrelated from

208    a statistical viewpoint (Creager 1971). We refer to this procedure as orthogonalization in keeping

209    with the conceptual goal of the analysis, although the term residualization is more statistically

210    accurate.

211

212    *Single-subject encoding*

213       Correlation distance (1 – Pearson's r) was used to compute the RDM of fMRI activity

214    patterns in each searchlight and each subject. Only voxels pertaining to the cerebral cortex with a

215    probability higher than 50% were included in the procedure (i.e., by applying a threshold over the

216    Harvard-Oxford cortical probabilistic atlas). After applying variance partitioning, the obtained $R^2$

217    for each component of unique and shared variance in each subject were z-scored and converted

218    into the partial correlation coefficient (de Heer et al. 2017) and then assigned to the center of the

219    searchlight, so obtaining a map for each subject and component. Results from four representative

220    subjects are shown in Figures S1-4, while the similarity between the full-model fits across subjects

221    is shown in Figure S5.

222

223    *Group-level test*

224       For each model, threshold free cluster enhancement (TFCE: Smith and Nichols 2009) was

225    used to detect group-level clusters significantly explained by the corresponding unique variance

226    component (5000 randomizations of the sign with 6mm variance smoothing, as implemented in

227    FSL's *randomise*: www.fmrib.ox.ac.uk/fsl/randomise). Statistical maps were then thresholded at

228    one-tailed $p < 0.05$, corrected for multiple comparisons across gray matter voxels and finally a

229    further arbitrary cluster size threshold of 10 voxels was applied (Figure 3).

230

10

*Orthogonality and complexity testing*

231

232    To look for differences in how information is encoded in different brain regions, we

233    introduced the orthogonality and complexity measures. Orthogonality was computed by dividing

234    the group-averaged (between subjects) sum of variance explained uniquely by the five models

235    with the group-averaged (between subjects) sum of variance explained by their shared

236    components for each searchlight (Orthogonality = $\Sigma$ $R^2_{\text{unique components}}$ / $\Sigma$ $R2_{\text{shared components}}$ ); a higher

237    value indicates, therefore, that a higher fraction of variance is explained by individual models,

238    rather than being shared across them. In principle, this value can take extreme values as 0, when

239    only shared information is coded, or infinite, when only orthogonal information is represented. To

240    evaluate whether orthogonality varies from posterior to anterior visual areas, we tested whether a

241    linear trend between the Y coordinate and average orthogonality within XZ-slices was present by

242    searching for abrupt changes in the slope, as high as 50% of the maximum value. As we found no

243    significant changes, the strength of the linear dependency between orthogonality and the

244    posterior-to-anterior direction was calculated using the Spearman's correlation (Figure 6A) and

245    significance was then computed with a parametric test.

246    Following (Vernon et al., 2016), two different groups of features were identified: low-level

247    representations, sensitive to retinotopic information, and abstract representations, that are

248    independent of the extent of retinotopic cortex stimulated. Inked-area and silhouette were

249    labeled as low-level models, and medial-axis, curvature and identity as abstract ones. Then,

250    complexity was measured by the ratio between variance uniquely explained either by low-level or

251    abstract models: Complexity = ($R^2_{\text{silhouette}}$ + $R^2_{\text{inked area}}$ ) / ($R^2_{\text{medial-axis}}$ + $R^2_{\text{curvature}}$ + $R^2_{\text{identity}}$). Thus,

252    within each searchlight, group-averaged (between subjects) sum of variance explained uniquely by

253    the low-level models was divided by the group-averaged sum of variance explained by the abstract

254    ones. Linearity was tested as for the orthogonality index.

11

255    Surface plots in Figures 4 and 5 were produced with the *Pycortex* toolbox for Python (Gao

256    et al. 2015). Second-level analyses were performed using custom-made code written in MATLAB

257    (MathWorks    Inc.).    Custom    Matlab    code    is    available    on    GitHub:

258    https://github.com/PPthe2nd/ShapeVar. Preprocessed data for all subjects and stimuli are publicly

259    available on Zenodo: https://doi.org/10.5281/zenodo.4038480. Access to the data is possible

260    upon submission of a Data User Agreement, available along with supplementary information on

261    Zenodo: https://doi.org/10.5281/zenodo.4003861.

262

263    **Results**

264    As expected from both theoretical and experimental investigations on this topic (Kay, 2011;

265    Bracci and Op de Beeck, 2016; Papale et al., 2019), the combination of our five models and our

266    stimulus set reveals moderate-to-high degrees of collinearity (Figure 2E). Consequently, to account

267    for multicollinearity before considering the significance of the association of each model with

268    brain representations, the variance partitioning analysis (Lescroart et al., 2015) and a searchlight

269    procedure (whole brain, 6mm radius: Kriegeskorte et al. 2006) were combined to identify group-

270    level clusters significantly explained by three shape models independently from competing

271    representations (Figure 2D).

272

273    *The human visual cortex encodes multiple orthogonal shape representations*

274    Group-level results show both distinct and overlapping clusters of shape selectivity in OTC,

275    also extending to posterior dorsal regions ($p < 0.05$ one-tailed, TFCE corrected). The silhouette

276    model (Figure 3, in red) shows a significant association with brain representations along the

277    Calcarine sulcus (CalcS), the occipitotemporal sulcus (OTS), the right collateral sulcus (CollS), the

278    right inferior temporal sulcus (ITS), the right fusiform gyrus (FusG), the cuneus (Cun) and in

12

279 posterior portions of the middle temporal gyrus (pMTG) and intraparietal sulcus (pIPS). The

280 medial-axis (Figure 3, in green) explains a significant portion of unique variance in the right lateral

281 occipital area (LO) only. Finally, curvature (Figure 3, in blue) significantly explains fMRI

282 representational geometries in the left lingual gyrus (LinG), in the bilateral FusG, along bilateral

283 OTS and ITS, along the right CollS, in the right MTG, bilaterally in the Cun and along the right IPS.

284 The significant clusters for the low- and high-level models are also represented in Figure 3.

285       As all orthogonal components of our tested models indicate the presence of at least a

286 significant cluster of selectivity, shape representation does not rely on a single feature, but on a

287 multi-dimensional coding scheme.

288

289           ** Figure 3 near here **

290

291 *Selectivity to orthogonal shape representations coexist in the same cortical regions*

292       We further explored the overlap between the selectivity to orthogonal shape

293 representations. Figure 4 depicts the pairwise comparisons between the three shape models in

294 our study. Qualitatively, a stronger overlap is observed in LO for medial-axis and curvature, and in

295 IT, right FusG, Cun, right pMTG and right pIPS for silhouette and curvature. Thus, those brain

296 regions encode multiple shape features, independently from the shared variance between them.

297 Also, both the effect sizes and spatial patterns are consistent across subjects (Figures S1-5).

298

299           ** Figure 4 near here **

300

301 *Topographic organization of object shape in right OTC*

13

302    Of note, all three models are significant only within right OTC (enclosed by a white line in

303    Figure 4). Figure 5 depicts right OTC in isolation with a greater detail: when combining the three

304    models (Figure 5B), a topographic organization emerges. Silhouette coding is medial with respect

305    to the CollS, encompassing the LinG and parahippocampal gyrus (PHG, red voxels in Figure 5B).

306    Proceeding laterally, the silhouette and medial-axis coexist in the fundus of the CollS (orange

307    voxels in Figure 5B), while the medial-axis extends also to the FusG (green voxels in Figure 5B).

308    Finally silhouette and curvature are both encoded medial to the OTS, with curvature being

309    encoded also in the fundus of the OTS.

310    These qualitative observations return a complex picture on shape coding in the human

311    brain. However, some general considerations can be made by looking at the interactions between

312    features.

313

314                                    ** Figure 5 near here **

315

316    *Coding of orthogonal object representations decreases from posterior to anterior regions*

317    In a previous study, Vernon *et al.* (2016) explored the relationship between retinotopic and

318    more abstract object representations, including contour curvature. They defined two orthogonal

319    components enclosing low-level and higher-level, complex features, and described a shift between

320    retinotopic and more abstract features in LO. Given the pattern of results from the previous

321    analysis, we further explored what has been observed by Vernon et al. (2016) and also tested the

322    relative weight of orthogonal and shared components. Indeed, the tuning to increasingly complex

323    features is considered a cornerstone of hierarchical object processing (Riesenhuber and Poggio

324    2000). However, it has been proposed that interaction between features may play a pivotal role in

325    evolving reliable selectivity in the brain (Benjamin et al. 2019). Thus, we hypothesized that shared

14

326  information should become more relevant along the visual hierarchy, moving from posterior to

327  anterior brain regions.

328     We defined two independent components, one for the low-level features and one for the

329  abstract ones, as in Vernon et al (2016). The first comprised the orthogonal variance of silhouette

330  and inked-area, since both are linked to the local retinotopic arrangement and to the extent of

331  retinotopic cortex stimulated. The second includes the orthogonal variance of medial-axis,

332  curvature (both insensitive to differences in object orientation and size) and object identity

333  models. The ratio between the explained variance of low-level and abstract features (i.e.,

334  complexity) in the posterior-to-anterior axis was computed by averaging the values in the XZ

335  plane: values higher than one indicate that brain representations are better accounted for by

336  retinotopic information, while values smaller than one indicate that abstract representations are

337  more relevant. When looking at the slope of complexity along the posterior-to-anterior axis, we

338  observed an abrupt shift (higher than 50% of the maximum) from retinotopic to abstract features

339  around $Y_{MNI}$ = -72 (Figure 6A). Of note, the shift occurs at the limit between occipital and temporal

340  or parietal cortex. As a matter of fact, previous studies (Haxby et al. 2001; Rice et al. 2014) on

341  ventral temporal cortex selectivity constrained their analysis between $Y_{MNI}$ = -70 and $Y_{MNI}$ = -20

342  (but see: Grill-Spector and Weiner 2014 for a different definition based on anatomical landmarks).

343     Then, we looked at the ratio between orthogonal and shared variance components (i.e.,

344  orthogonality) in the posterior-to-anterior axis. The variance explained by the orthogonal

345  components of the five models was first summed, and then divided by the sum of the shared

346  components between the five models. Here, values higher than one indicate that brain

347  representations are better explained by orthogonal components of variance. Orthogonality

348  linearly decreases along the posterior-to-anterior axis, without shifts ($\rho$ = -0.83, p < 0.001,

349  parametric test; Figure 6B). Thus, while orthogonal information is always more represented than

15

350    shared variance (min = 2.15), it becomes less relevant proceeding along the visual hierarchy, and

351    neither ventral nor dorsal streams alone are responsible for the overall effect (Figure S6).

352        Of note, both these ratios abstract away from magnitude of response and goodness of fit

353    *per se*. Instead, they reveal the relative contribution of different features to different brain areas

354    beyond their overall performance.

355

356                                              ** Figure 6 near here **

357

358    **Discussion**

359        In the present study, we found that object shape is not represented by a single feature but

360    is encoded by multiple representations (i.e., silhouette, medial-axis and curvature) that uniquely

361    contribute to object processing in the human visual cortex (Figures 3-5). Moreover, we showed

362    that the brain encodes orthogonal object representations in a topographic fashion: the early visual

363    cortex is biased towards unique components of variance, while shared representations become

364    progressively more relevant in more anterior regions (Figure 6).

365

366    *Shape coding is multidimensional*

367        In line with previous studies, we found that object silhouette is mainly encoded in early

368    visual areas (Bracci and Op de Beeck 2016; Kaiser et al. 2016; Khaligh-Razavi and Kriegeskorte

369    2014; Proklova et al. 2016). This result can be explained by top-down figure-dependent

370    mechanisms that modulate V1 activity both in monkeys (Poort et al. 2016; Self et al. 2019) and

371    humans (Kok and de Lange 2014; Muckli et al. 2015), and enhances the processing of object-

372    related information in early visual areas also during natural vision (Papale et al. 2018). However,

16

373 another possibility may be that the silhouette model better captures the object physical

374 appearance (Kubilius et al. 2016).

375 Instead, the variance component unique to the medial-axis model – which is the most

376 transformation-resistant shape description (Yang et al. 2008) – was significant in a smaller extent

377 of cortex comprising only a subset of voxels in right LO (Figure 3, middle in green). This can be due

378 to a higher spatial inter-subject variability of this representation that has been already observed

379 by Leeds *et al.* (2013), or to a higher collinearity with the low- and high-level models we employed

380 (Figure 2D) that prevents from disentangling its contribution from competing representations.

381 Nonetheless, our result fits previous evidence of medial-axis coding in monkey IT (Hung et al.

382 2012; putative homologue of human LO), human LO (while also controlling for low-level

383 properties: Ayzenberg et al. 2019b) and is consistent with our previous MEG study showing that

384 medial-axis processing is limited to a small cluster of right posterior sensors, when controlling for

385 collinearity with low-level and categorical representations (Papale et al. 2019).

386 Finally, IT (Kayaert et al. 2005b; Yue et al. 2014), LO (Vernon et al. 2016) and FusG (Caldara

387 et al. 2006) were bilaterally tuned to contour curvature (Figure 3, bottom in blue), in accordance

388 with previous neuroimaging investigations. Actually, LO has a pivotal role in object processing

389 (Grill-Spector et al. 2001; Grill-Spector et al. 1999; Kourtzi and Kanwisher 2001), as IT in monkeys

390 (Brincat and Connor 2004; Desimone et al. 1984; Kayaert et al. 2005a; Op de Beeck et al. 2001;

391 Tanaka 2003; Zoccolan et al. 2007). In addition, while we focus our discussion on the ventral

392 stream, we also observed few significant clusters in dorsal visual regions (R pIPS; see Figure 3),

393 both for curvature and silhouette, which confirm previous observations (Freud et al. 2017).

394 Our result favors the hypothesis of a key role of right OTC in coding object shape. In line

395 with our view, recent neuropsychological evidence found visual agnosia in a subject with a cortical

396 lesion circumscribed to right OTC. Moreover, this focal right OTC lesion affected shape selectivity

17

397    across the whole visual cortex, leading to large-scale alterations that were stable over time (Freud

398    and Behrmann 2020).

399    Overall, the evidence that all the tested dimensions independently contribute to shape

400    representation in the human visual cortex favors the hypothesis of a multi-dimensional coding of

401    object shape (Silson et al. 2016; Silson et al. 2013), similarly to what is observed for texture

402    processing (Okazawa et al. 2015; Ziemba et al. 2016). In line with this, a recent study showed that

403    behavioral shape similarity could be modeled with no less than 109 different dimensions

404    (Morgenstern et al. 2020).

405

406    *Coding of shared information increases along the visual hierarchy*

407    Long *et al.* (2018) suggested that mid-level computations, covarying with high-level semantic

408    processing (including curvature extraction), control the organization of OTC. In the present study,

409    however, we observed overlapping selectivity to orthogonal features in LO (medial-axis and

410    curvature), IT, right FusG, Cun, right pMTG and right pIPS (silhouette and curvature). Since we

411    controlled for collinearity between models, this result could not be merely ascribed to the

412    variance shared by those features. Here, we also observed that coding of shared descriptions in

413    OTC is topographically arranged and its relevance linearly increases from posterior to anterior

414    regions (Figure 6). This observation, consistent with the core finding of Long *et al.* (2018), suggests

415    that the hierarchy of visual processing is not only shaped by specificity to increasingly complex

416    features, but also by a higher selectivity to shared representations.

417    This observation complements what has been already observed on the two extremes of

418    the ventral visual pathway: V1 and IT. Representations in V1 are over-complete relative to the

419    retinal input (Olshausen and Field 1996; Vinje and Gallant 2000). In addition, inhibitory

420    interactions in V1 are specifically targeted at neurons with similar tuning properties (Chettih and

18

421   Harvey 2019). Both these factors increase V1 representational capacity and may ultimately lead to

422   a higher selectivity to orthogonal features, as we observed in posterior regions. On the other

423   hand, higher sensitivity to shared information in more anterior areas may be produced by

424   populations of neurons that are not tuned to a specific property but that encode multiple

425   dimensions at once. Indeed, shared featural selectivity has been proposed as the mechanism

426   responsible to achieve dimensionality reduction of the sensory input in IT (Lehky et al. 2014),

427   where both neural density and surface are much lower than in V1 (Cahalane et al. 2012; Van Essen

428   et al. 1992). In line with this, the highest-dimensional among our three shape models (i.e.,

429   silhouette) is also represented in posterior regions (Figure 3). Relatedly, the interaction between

430   multiple features is thought to represent the optimal solution to increase the sensitivity to their

431   mutual changes: in this view, instead of having few neurons encoding a single feature each, it may

432   be preferable to have most of the neurons encoding multiple features at once (Benjamin et al.

433   2019). It has been also suggested that interactions between features are responsible for the poor

434   reliability of tuning curves in predicting brain responses in natural vision (Benjamin et al. 2019).

435         Thus, what can be concluded on the nature of object processing? On one hand, we

436   observed an abrupt shift from retinotopic to abstract representations moving anteriorly across the

437   brain (Figure 6A). However, this shift is relative: though less relevant, orthogonal retinotopic

438   information spreads also to OTC, explaining a significant portion of its variance, in line with

439   previous work and suggesting a link between low-level and object selectivity (Rajimehr et al. 2011;

440   Rice et al. 2014). On the other hand, we found a linear dependency between the posterior-to-

441   anterior axis and the variance explained by shared information (Figure 6B), in line with previous

442   research showing an increase in coding category-orthogonal information from V4 to IT in non-

443   human primates (Hong et al. 2016). As stated earlier, this property describes the linear cascade of

444   computations in the visual hierarchy better than complexity: optimizing the coding of shared

19

445  variance between behaviorally relevant features may represent a key factor in shaping the

446  architecture of our visual cortex and achieving reliable, view-point invariant object

447  representations. In this light, the next step should be to move from modeling representational

448  geometries to more direct modulations of brain responses, so to control also for nonlinear

449  interactions between features (Benjamin et al. 2019).

450

451  *Limits and conclusions*

452  It should be noted that due to the low fMRI temporal resolution, we cannot resolve which

453  mechanisms support the different tuning for shared representations. Moreover, while the

454  selected models capture visual transformations, many alternative descriptions exist (e.g., Khaligh-

455  Razavi and Kriegeskorte 2014). In addition, the searchlight procedure introduces some

456  unavoidable imprecision in localization (but see: Lettieri et al. 2019 for an analytical exploration of

457  this issue), thus further studies using the same shape descriptions might find slightly different

458  locations as those found in this study. Overall, however, our results hint at the existence of a

459  multi-dimensional coding of object shape, and reveal that selectivity for shared object

460  representations are topographically arranged and increases along the visual hierarchy. Future

461  experiments will identify how different tasks (e.g., determining object similarity vs. extracting

462  affordances), and alternative descriptions impact on the observed patterns of selectivity. Finally,

463  we described what qualitatively appears to be a shape coding topography in right OTC: further

464  research would be necessary to understand how strong is the link between structural and

465  functional organizations.

466

467  **Acknowledgments**

20

469 grants PRIN 2015WXAXJF and 2015AR52F9.

470

471 **Author contributions**

472 Pa.P., A.L. G.H., L.C. and E.R. conceived the study. A.L., G.H. and L.C. performed experiments.

473 Pa.P., and A.L. analyzed the data. All the authors discussed the results and wrote the manuscript.

474

475 **References**

476

477 **Ayzenberg V, Chen Y, Yousif SR, and Lourenco SF**. Skeletal representations of shape in human

478 vision: Evidence for a pruned medial axis model. *J Vis* 19: 6, 2019a.

479 **Ayzenberg V, Kamps FS, Dilks DD, and Lourenco SF**. A dual role for shape skeletons in human

480 vision: perceptual organization and object recognition. *bioRxiv* 799650, 2019b.

481 **Ayzenberg V, and Lourenco SF**. Skeletal descriptions of shape provide unique perceptual

482 information for object recognition. *Sci Rep* 9: 9359, 2019.

483 **Benjamin AS, Ramkumar P, Fernandes H, Smith MA, and Kording KP**. Hue tuning curves in V4

484 change with visual context. *bioRxiv* 780478, 2019.

485 **Blum H**. Biological shape and visual science. I. *J Theor Biol* 38: 205-287, 1973.

486 **Bracci S, and Op de Beeck H**. Dissociations and Associations between Shape and Category

487 Representations in the Two Visual Pathways. *J Neurosci* 36: 432-444, 2016.

21

488   **Bracci S, Ritchie JB, Kalfas I, and Op de Beeck H**. The ventral visual pathway represents animal

489   appearance over animacy, unlike human behavior and deep neural networks. *J Neurosci* 2019.


490   **Brincat SL, and Connor CE**. Underlying principles of visual shape selectivity in posterior

491   inferotemporal cortex. *Nat Neurosci* 7: 880-886, 2004.


492   **Cadieu C, Kouh M, Pasupathy A, Connor CE, Riesenhuber M, and Poggio T**. A model of V4 shape

493   selectivity and invariance. *J Neurophysiol* 98: 1733-1750, 2007.


494   **Cahalane DJ, Charvet CJ, and Finlay BL**. Systematic, balancing gradients in neuron density and

495   number across the primate isocortex. *Front Neuroanat* 6: 28, 2012.


496   **Caldara R, Seghier ML, Rossion B, Lazeyras F, Michel C, and Hauert CA**. The fusiform face area is

497   tuned for curvilinear patterns with more high-contrasted elements in the upper part. *Neuroimage*

498   31: 313-319, 2006.


499   **Carlson ET, Rasquinha RJ, Zhang K, and Connor CE**. A sparse object coding scheme in area V4. *Curr*

500   *Biol* 21: 288-293, 2011.


501   **Chettih SN, and Harvey CD**. Single-neuron perturbations reveal feature-specific competition in V1.

502   *Nature* 567: 334-340, 2019.


503   **Connor CE, Brincat SL, and Pasupathy A**. Transformation of shape information in the ventral

504   pathway. *Curr Opin Neurobiol* 17: 140-147, 2007.

505    **Cox RW**. AFNI: software for analysis and visualization of functional magnetic resonance

506    neuroimages. *Comput Biomed Res* 29: 162-173, 1996.


507    **Creager JA**. Comments: Orthogonal and Nonorthogonal Methods for Partitioning Regression

508    Variance. *American Educational Research Journal* 8: 671-676, 1971.


509    **de Heer WA, Huth AG, Griffiths TL, Gallant JL, and Theunissen FE**. The Hierarchical Cortical

510    Organization of Human Speech Processing. *J Neurosci* 37: 6539-6557, 2017.


511    **Desimone R, Albright TD, Gross CG, and Bruce C**. Stimulus-selective properties of inferior

512    temporal neurons in the macaque. *J Neurosci* 4: 2051-2062, 1984.


513    **Elder JH, and Velisavljevic L**. Cue dynamics underlying rapid detection of animals in natural

514    scenes. *J Vis* 9: 7, 2009.


515    **Freud E, and Behrmann M**. Altered large-scale organization of shape processing in visual agnosia.

516    *Cortex* 2020.


517    **Freud E, Culham JC, Plaut DC, and Behrmann M**. The large-scale organization of shape processing

518    in the ventral and dorsal pathways. *Elife* 6: 2017.


519    **Gao JS, Huth AG, Lescroart MD, and Gallant JL**. Pycortex: an interactive surface visualizer for fMRI.

520    *Front Neuroinform* 9: 23, 2015.


521    **Greene MR, Baldassano C, Esteva A, Beck DM, and Fei-Fei L**. Visual scenes are categorized by

522    function. *J Exp Psychol Gen* 145: 82-94, 2016.

23

523    **Grill-Spector K, Kourtzi Z, and Kanwisher N**. The lateral occipital complex and its role in object

524    recognition. *Vision Res* 41: 1409-1422, 2001.


525    **Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, and Malach R**. Differential processing

526    of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24:

527    187-203, 1999.


528    **Grill-Spector K, and Weiner KS**. The functional architecture of the ventral temporal cortex and its

529    role in categorization. *Nature Reviews Neuroscience* 15: 536-548, 2014.


530    **Groen, II, Ghebreab S, Lamme VA, and Scholte HS**. Spatially pooled contrast responses predict

531    neural and perceptual similarity of naturalistic image categories. *PLOS Comput Biol* 8: e1002726,

532    2012.


533    **Groen, II, Greene MR, Baldassano C, Fei-Fei L, Beck DM, and Baker CI**. Distinct contributions of

534    functional and deep neural network features to representational similarity of scenes in human

535    brain and behavior. *Elife* 7: 2018.


536    **Handjaras G, Leo A, Cecchetti L, Papale P, Lenci A, Marotta G, Pietrini P, and Ricciardi E**.

537    Modality-independent encoding of individual concepts in the left parietal cortex.

538    *Neuropsychologia* 105: 39-49, 2017.


539    **Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, and Pietrini P**. Distributed and overlapping

540    representations of faces and objects in ventral temporal cortex. *Science* 293: 2425-2430, 2001.

24

541    **Hebart MN, Bankson BB, Harel A, Baker CI, and Cichy RM**. The representational dynamics of task

542    and object processing in humans. *eLife* 7: e32816, 2018.

543    **Hong H, Yamins DL, Majaj NJ, and DiCarlo JJ**. Explicit information for category-orthogonal object

544    properties increases along the ventral stream. *Nat Neurosci* 19: 613-622, 2016.

545    **Hung CC, Carlson ET, and Connor CE**. Medial axis shape coding in macaque inferotemporal cortex.

546    *Neuron* 74: 1099-1113, 2012.

547    **Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, and Smith SM**. Fsl. *Neuroimage* 62: 782-

548    790, 2012.

549    **Kaiser D, Azzalini DC, and Peelen MV**. Shape-independent object category responses revealed by

550    MEG and fMRI decoding. *J Neurophysiol* 115: 2246-2250, 2016.

551    **Kay KN**. Understanding visual representation by developing receptive-field models. *Visual*

552    *population codes: Towards a common multivariate framework for cell recording and functional*

553    *imaging* 133-162, 2011.

554    **Kayaert G, Biederman I, Op de Beeck HP, and Vogels R**. Tuning for shape dimensions in macaque

555    inferior temporal cortex. *Eur J Neurosci* 22: 212-224, 2005a.

556    **Kayaert G, Biederman I, and Vogels R**. Representation of regular and irregular shapes in macaque

557    inferotemporal cortex. *Cereb Cortex* 15: 1308-1321, 2005b.

558 **Khaligh-Razavi S-M, and Kriegeskorte N**. Deep supervised, but not unsupervised, models may

559 explain IT cortical representation. 2014.

560 **Kok P, and de Lange FP**. Shape perception simultaneously up- and downregulates neural activity in

561 the primary visual cortex. *Curr Biol* 24: 1531-1535, 2014.

562 **Kourtzi Z, and Kanwisher N**. Representation of perceived object shape by the human lateral

563 occipital complex. *Science* 293: 1506-1509, 2001.

564 **Kriegeskorte N, Goebel R, and Bandettini P**. Information-based functional brain mapping.

565 *Proceedings of the National Academy of Sciences of the United States of America* 103: 3863-3868,

566 2006.

567 **Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, and Bandettini PA**.

568 Matching categorical object representations in inferior temporal cortex of man and monkey.

569 *Neuron* 60: 1126-1141, 2008.

570 **Kubilius J, Bracci S, and Op de Beeck HP**. Deep Neural Networks as a Computational Model for

571 Human Shape Sensitivity. *PLOS Comput Biol* 12: e1004896, 2016.

572 **Lawrence SJ, Keefe BD, Vernon RJ, Wade AR, McKeefry DJ, and Morland AB**. Global shape

573 aftereffects in composite radial frequency patterns. *J Vis* 16: 17, 2016.

574 **Leeds DD, Seibert DA, Pyles JA, and Tarr MJ**. Comparing visual representations across human fMRI

575 and computational vision. *J Vis* 13: 25, 2013.

576 **Lehky SR, Kiani R, Esteky H, and Tanaka K**. Dimensionality of object representations in monkey

577 inferotemporal cortex. *Neural Comput* 26: 2135-2162, 2014.


578 **Lescroart MD, and Biederman I**. Cortical representation of medial axis structure. *Cereb Cortex* 23:

579 629-637, 2013.


580 **Lescroart MD, Stansbury DE, and Gallant JL**. Fourier power, subjective distance, and object

581 categories all provide plausible models of BOLD responses in scene-selective visual areas. *Front*

582 *Comput Neurosci* 9: 135, 2015.


583 **Lettieri G, Handjaras G, Ricciardi E, Leo A, Papale P, Betta M, Pietrini P, and Cecchetti L**.

584 Emotionotopy in the human right temporo-parietal cortex. *Nat Commun* 10: 5568, 2019.


585 **Long B, Stormer VS, and Alvarez GA**. Mid-level perceptual features contain early cues to animacy.

586 *J Vis* 17: 20, 2017.


587 **Long B, Yu CP, and Konkle T**. Mid-level visual features underlie the high-level categorical

588 organization of the ventral stream. *Proc Natl Acad Sci U S A* 115: E9015-E9024, 2018.


589 **Lowet AS, Firestone C, and Scholl BJ**. Seeing structure: Shape skeletons modulate perceived

590 similarity. *Atten Percept Psychophys* 80: 1278-1289, 2018.


591 **Monroy A, Eigenstetter A, and Ommer B**. Beyond straight lines—object detection using curvature.

592 IEEE, 2011, p. 3561-3564.

593    **Morgenstern Y, Hartmann F, Schmidt F, Tiedemann H, Prokott E, Maiello G, and Fleming RW**. An

594    image-computable model of human visual shape similarity. *bioRxiv* 2020.


595    **Muckli L, De Martino F, Vizioli L, Petro LS, Smith FW, Ugurbil K, Goebel R, and Yacoub E**.

596    Contextual Feedback to Superficial Layers of V1. *Curr Biol* 25: 2690-2695, 2015.


597    **Nimon K, Lewis M, Kane R, and Haynes RM**. An R package to compute commonality coefficients in

598    the multiple regression case: an introduction to the package and a practical example. *Behav Res*

599    *Methods* 40: 457-466, 2008.


600    **Nimon KF, and Oswald FL**. Understanding the results of multiple linear regression: Beyond

601    standardized regression coefficients. *Organizational Research Methods* 16: 650-674, 2013.


602    **Okazawa G, Tajima S, and Komatsu H**. Image statistics underlying natural texture selectivity of

603    neurons in macaque V4. *Proc Natl Acad Sci U S A* 112: E351-360, 2015.


604    **Olshausen BA, and Field DJ**. Emergence of simple-cell receptive field properties by learning a

605    sparse code for natural images. *Nature* 381: 607-609, 1996.


606    **Op de Beeck H, Wagemans J, and Vogels R**. Inferotemporal neurons represent low-dimensional

607    configurations of parameterized shapes. *Nat Neurosci* 4: 1244-1252, 2001.


608    **Papale P, Betta M, Handjaras G, Malfatti G, Cecchetti L, Rampinini A, Pietrini P, Ricciardi E,**

609    **Turella L, and Leo A**. Common spatiotemporal processing of visual features shapes object

610    representation. *Sci Rep* 9: 7601, 2019.

28

611 **Papale P, Leo A, Cecchetti L, Handjaras G, Kay KN, Pietrini P, and Ricciardi E**. Foreground-

612 Background Segmentation Revealed during Natural Image Viewing. *eNeuro* 5: 2018.

613 **Poort J, Self MW, van Vugt B, Malkki H, and Roelfsema PR**. Texture Segregation Causes Early

614 Figure Enhancement and Later Ground Suppression in Areas V1 and V4 of Visual Cortex. *Cereb*

615 *Cortex* 26: 3964-3976, 2016.

616 **Proklova D, Kaiser D, and Peelen MV**. Disentangling Representations of Object Shape and Object

617 Category in Human Visual Cortex: The Animate-Inanimate Distinction. *J Cogn Neurosci* 28: 680-

618 692, 2016.

619 **Rajimehr R, Devaney KJ, Bilenko NY, Young JC, and Tootell RBH**. The "Parahippocampal Place

620 Area" Responds Preferentially to High Spatial Frequencies in Humans and Monkeys. *PLoS Biol* 9:

621 e1000608, 2011.

622 **Ramakrishnan K, Scholte HS, Groen, II, Smeulders AW, and Ghebreab S**. Visual dictionaries as

623 intermediate features in the human brain. *Front Comput Neurosci* 8: 168, 2014.

624 **Rice GE, Watson DM, Hartley T, and Andrews TJ**. Low-level image properties of visual objects

625 predict patterns of neural response across category-selective regions of the ventral visual

626 pathway. *The Journal of Neuroscience* 34: 8837-8844, 2014.

627 **Riesenhuber M, and Poggio T**. Models of object recognition. *Nature neuroscience* 3: 1199-1204,

628 2000.

629    **Sebastian TB, Klein PN, and Kimia BB**. Recognition of shapes by editing their shock graphs. *IEEE*

630    *Trans Pattern Anal Mach Intell* 26: 550-571, 2004.

631    **Self MW, Jeurissen D, van Ham AF, van Vugt B, Poort J, and Roelfsema PR**. The Segmentation of

632    Proto-Objects in the Monkey Primary Visual Cortex. *Curr Biol* 29: 1019-1029 e1014, 2019.

633    **Silson EH, Groen, II, Kravitz DJ, and Baker CI**. Evaluating the correspondence between face-,

634    scene-, and object-selectivity and retinotopic organization within lateral occipitotemporal cortex. *J*

635    *Vis* 16: 14, 2016.

636    **Silson EH, McKeefry DJ, Rodgers J, Gouws AD, Hymers M, and Morland AB**. Specialized and

637    independent processing of orientation and shape in visual field maps LO1 and LO2. *Nat Neurosci*

638    16: 267-269, 2013.

639    **Smith SM, and Nichols TE**. Threshold-free cluster enhancement: addressing problems of

640    smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44: 83-98,

641    2009.

642    **Tanaka K**. Columns for complex visual object features in the inferotemporal cortex: clustering of

643    cells with similar but slightly different stimulus selectivities. *Cereb Cortex* 13: 90-99, 2003.

644    **Van Eede M, Macrini D, Telea A, Sminchisescu C, and Dickinson SS**. Canonical skeletons for shape

645    matching. IEEE, 2006, p. 64-69.

646    **Van Essen DC, Anderson CH, and Felleman DJ**. Information processing in the primate visual

647    system: an integrated systems perspective. *Science* 255: 419-423, 1992.

30

648  **Vernon RJ, Gouws AD, Lawrence SJ, Wade AR, and Morland AB**. Multivariate Patterns in the

649  Human Object-Processing Pathway Reveal a Shift from Retinotopic to Shape Curvature

650  Representations in Lateral Occipital Areas, LO-1 and LO-2. *J Neurosci* 36: 5763-5774, 2016.


651  **Vinje WE, and Gallant JL**. Sparse coding and decorrelation in primary visual cortex during natural

652  vision. *Science* 287: 1273-1276, 2000.


653  **Wolfe JM, Yee A, and Friedman-Hill SR**. Curvature is a basic feature for visual search tasks.

654  *Perception* 21: 465-480, 1992.


655  **Yang M, Kpalma K, and Ronsin J**. *A survey of shape feature extraction techniques*. In-Tech, 2008.


656  **Yue X, Pourladian IS, Tootell RB, and Ungerleider LG**. Curvature-processing network in macaque

657  visual cortex. *Proc Natl Acad Sci U S A* 111: E3467-3475, 2014.


658  **Ziemba CM, Freeman J, Movshon JA, and Simoncelli EP**. Selectivity and tolerance for visual

659  texture in macaque V2. *Proc Natl Acad Sci U S A* 113: E3140-3149, 2016.


660  **Zoccolan D, Kouh M, Poggio T, and DiCarlo JJ**. Trade-off between object selectivity and tolerance

661  in monkey inferotemporal cortex. *J Neurosci* 27: 12292-12307, 2007.

662
663

31

664 **Figure 1**. **How does our brain encode object shape?**

665 Different features capture specific aspects of object shape. For instance, silhouette and curvature

666 descriptions of the same shapes may be orthogonal to each other (red- and blue-shaded areas) or

667 vary in a linear fashion (purple-shaded area). Thus, our brain may represent object shape by

668 extracting one specific and more reliable feature, by focusing on shared representations across

669 multiple features, or even encoding the orthogonal components of different features.

670

671 **Figure 2**. **Schematic of the shape models and experiment.**

672 A) Five different object representations are employed: three shape models and two further

673 descriptions. From left: silhouette, medial axis, curvature, inked area (low-level) and object

674 identity (high-level).

675 B) Representational dissimilarity matrices (RDMs) of each model: they represent all the possible

676 pairwise distances between the stimuli.

677 C) Methodological pipeline. Brain responses were recorded while subjects maintained fixation on

678 a colored fixation cross, paying attention to color switching between red and green. Orthogonal to

679 the task, we presented 42 grayscale pictures of real objects, for 1s each. Activity patterns were

680 used to test the association between the five model RDMs and each brain activity RDM, computed

681 combining a searchlight procedure with a variance partitioning analysis.

682 D) Similarity between the five model RDMs. As expected, the five representations are correlated.

683 However, the variance partitioning approach controls for the effect of model collinearity.

684

685

686 **Figure 3. The human visual cortex encodes orthogonal shape representations.**

687    Group-level maps showing significant clusters of shape selectivity in OTC and in posterior dorsal

688    regions (one-tailed p < 0.05, TFCE corrected). Each row of images depicts selectivity to orthogonal

689    components of curvature (blue), silhouette (red), medial-axis (green), inked area (purple) and

690    object identity (orange). CalcS: Calcarine sulcus; OTS: occipitotemporal sulcus; CollS: collateral

691    sulcus; ITS: inferior temporal sulcus; FusG: fusiform gyrus; Cun: cuneus; MTG: middle temporal

692    gyrus; IPS: intraparietal sulcus; IPC: intraparietal cortex.

693

694    **Figure 4. Coding of orthogonal shape components overlap in the human visual cortex**

695    Pairwise comparisons between group-level unthresholded T-maps of orthogonal shape

696    components show that several regions encode more than a single orthogonal description. Colored

697    voxels have high T-value in a single model. Silhouette is represented in red, medial-axis in green

698    and curvature in blue. The overlap between two orthogonal representations is indicated by white

699    voxels, while brightness represents the value of the T-statistic in each voxel (i.e. gray and black

700    voxels have low T-value in both models). White lines enclose right OTC, where all three shape

701    models are significant.

702

703    **Figure 5. Topographic organization of object shape in right OTC**

704    A) Pairwise comparisons between group-level T-maps of orthogonal shape components in right

705        OTC. Colored voxels have high T-value in a single model. Silhouette is represented in red,

706        medial-axis in green and curvature in blue. The overlap between two orthogonal

707        representations is indicated by white voxels, while brightness represents the value of the T-

708        statistic in each voxel (i.e. gray and black voxels have low T-value in both models).

709    B) Overlap between the three group-level T-maps of orthogonal shape components in right OTC.

710        Silhouette is represented in red, medial-axis in green and curvature in blue. The overlap
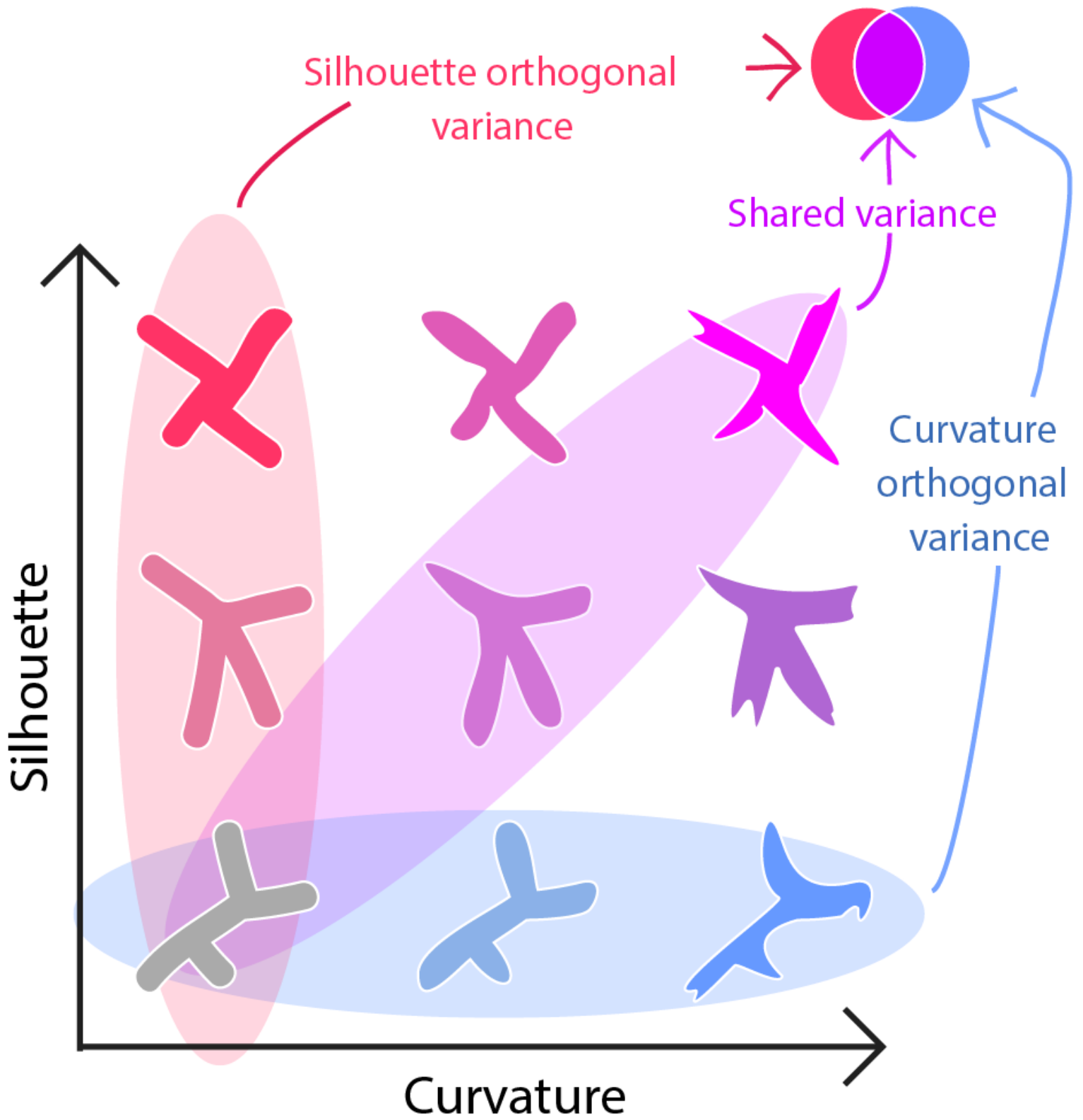
33

711    between two orthogonal representations is indicated by intermediate colors: pink for

712    silhouette and curvature, orange for silhouette and medial-axis, cyan for medial-axis and

713    curvature. Brightness represents the value of the T-statistic in each voxel (i.e. gray and black

714    voxels have low T-value in all models). The solid lines show the borders of the four slices

715    depicted in (C): these are arranged from more posterior (top) to more anterior (bottom) and

716    are enclosed between points (A-B and A'-B') referring to their medial-to-lateral extent.
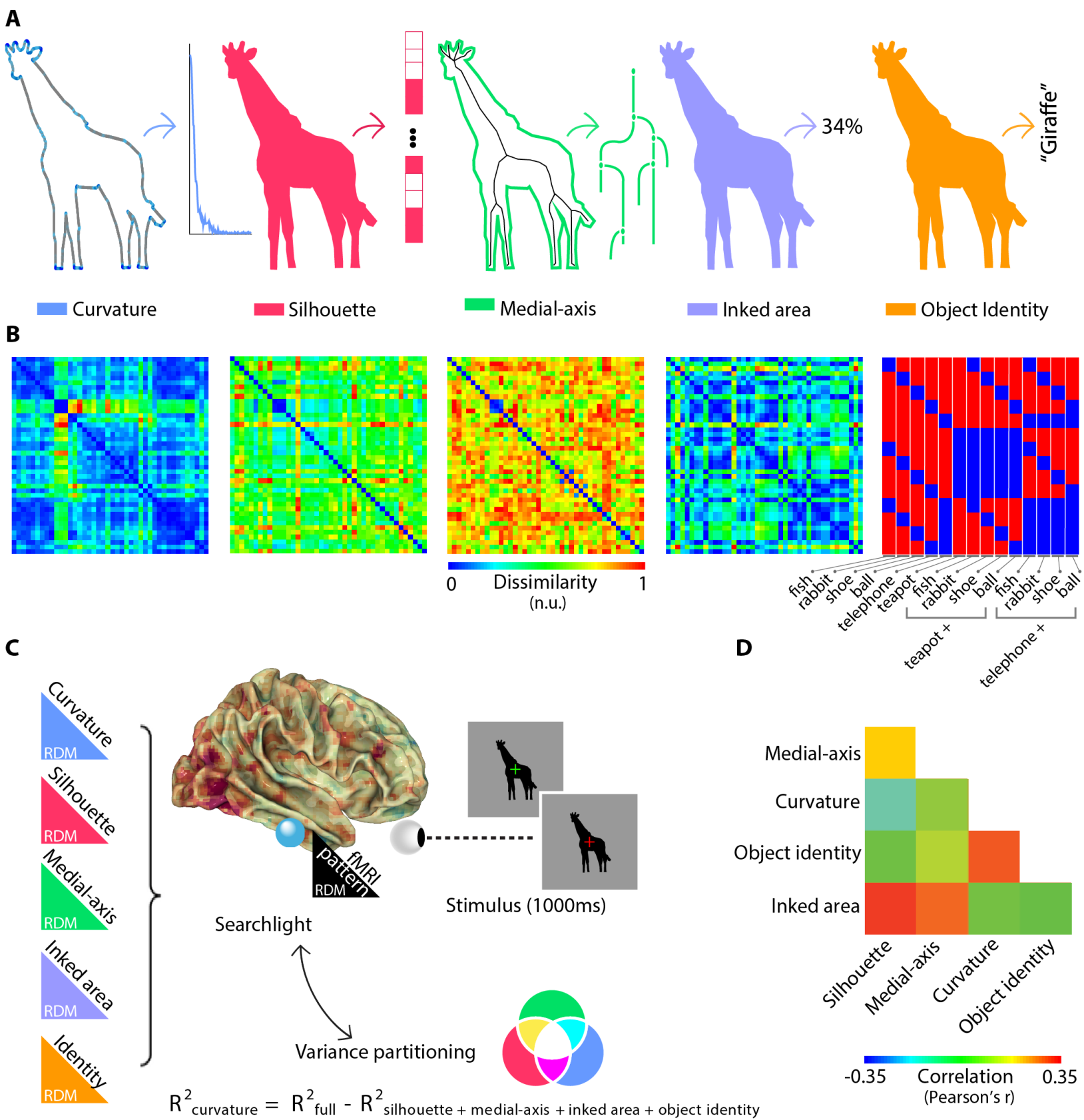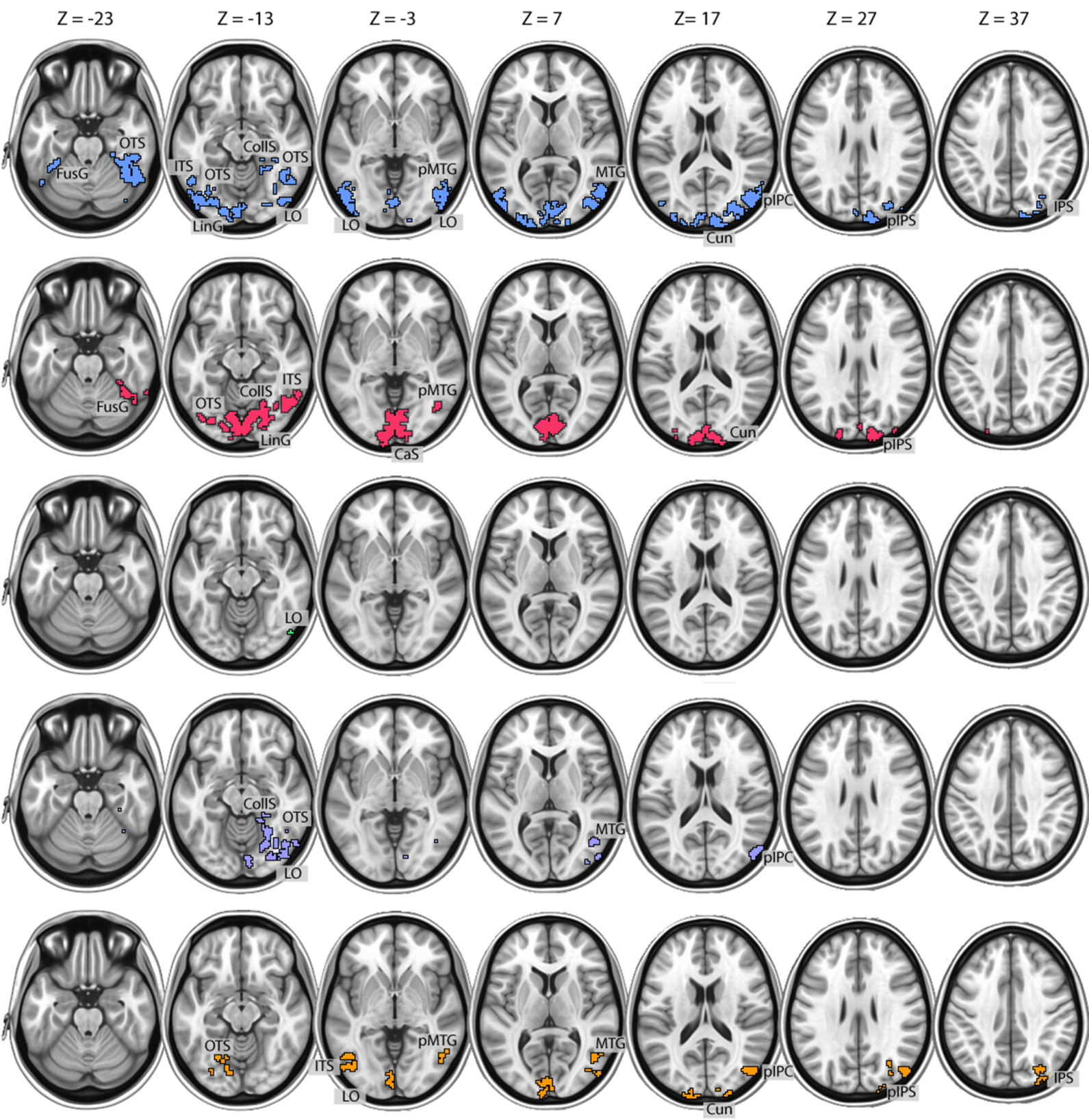
717

718    **Figure 6. The link between object features shapes the human visual hierarchy**

719    A) The ratio between the explained variance of low-level and abstract features (i.e. complexity)

720    along the visual hierarchy reveals an abrupt shift. Values higher than one (horizontal dashed line)

721    indicate that brain representations are better accounted for by retinotopic information while

722    values smaller than one that abstract representation is more relevant. The vertical dashed line

723    represents the point where mean and slope (dashed black lines) present an abrupt change.
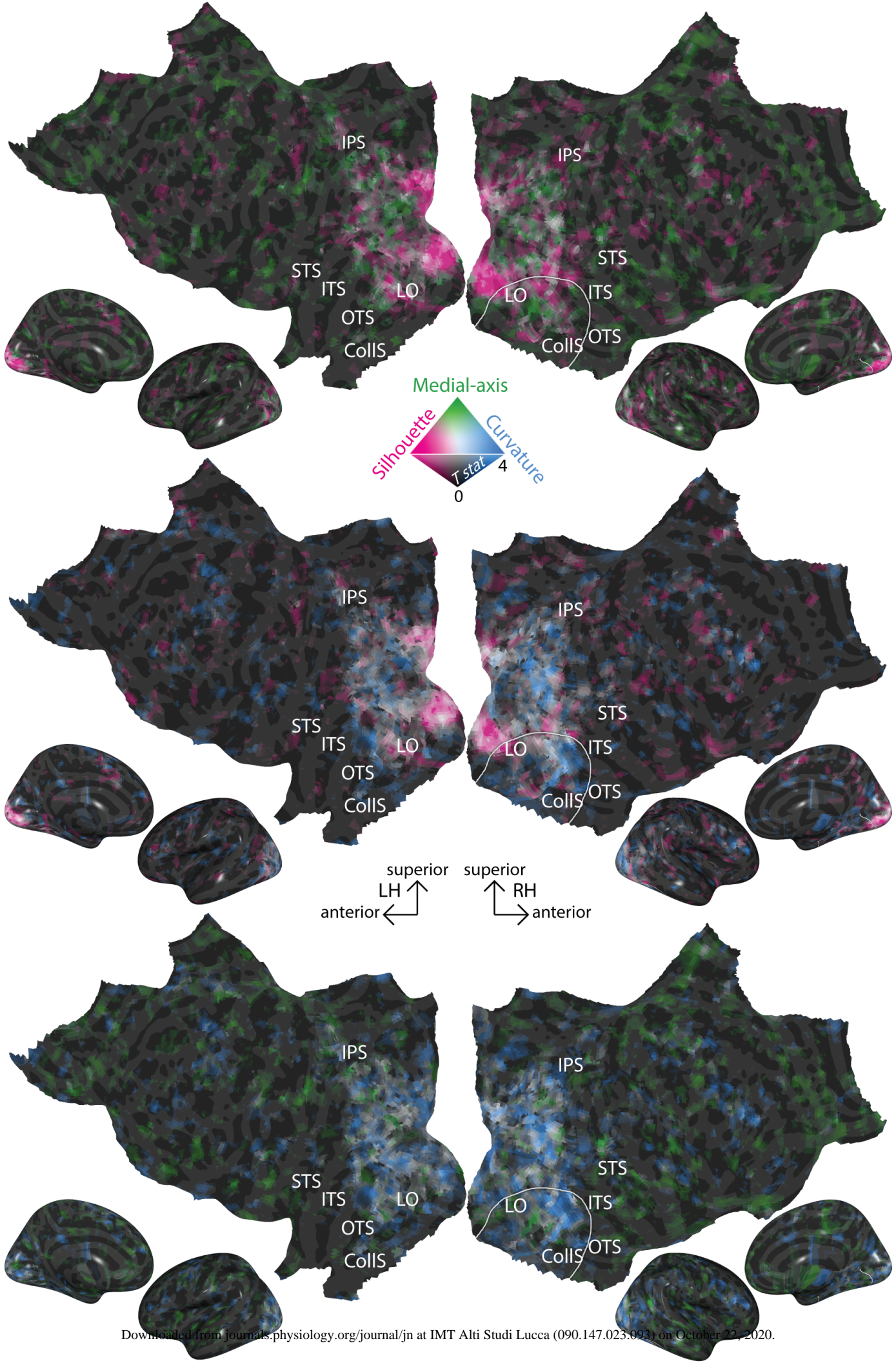
724    B) The ratio between the variance explained by the orthogonal components of the five models and

725    the sum of the shared components between the five models (i.e. orthogonality) linearly decreases

726    along the visual hierarchy ($\rho$ = -0.83, ***: p < 0.001, parametric test). Values higher than one

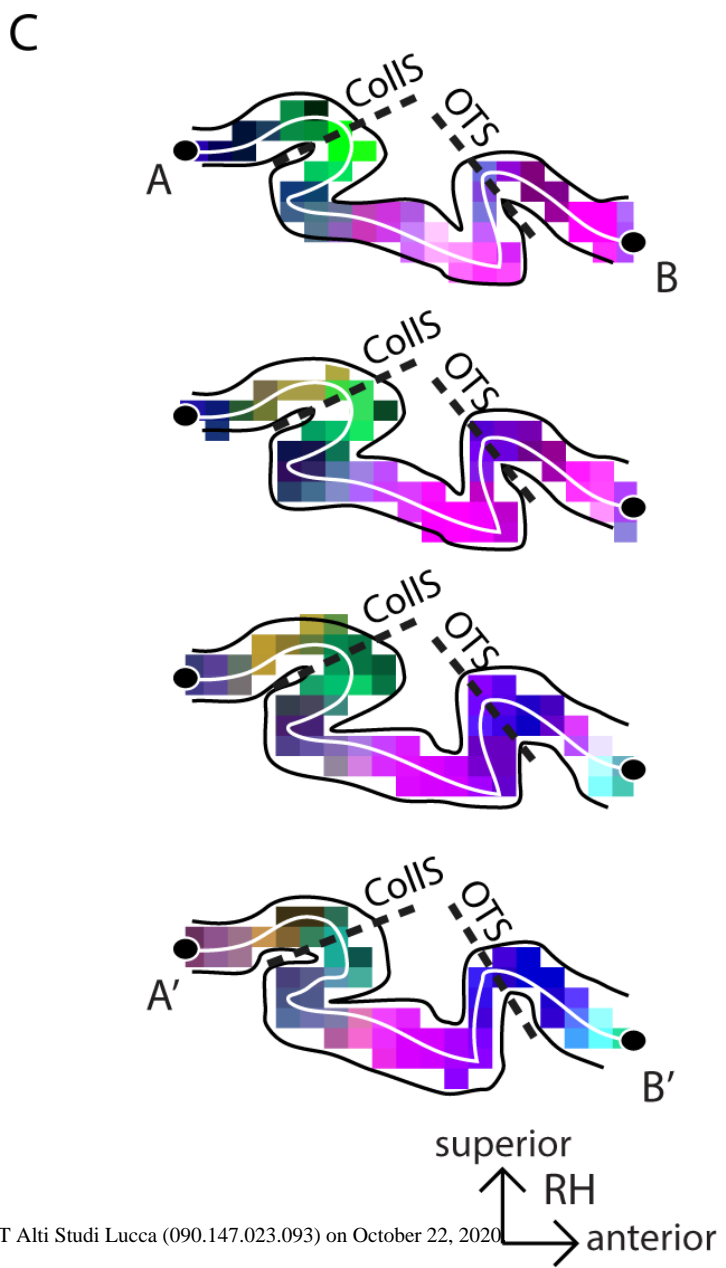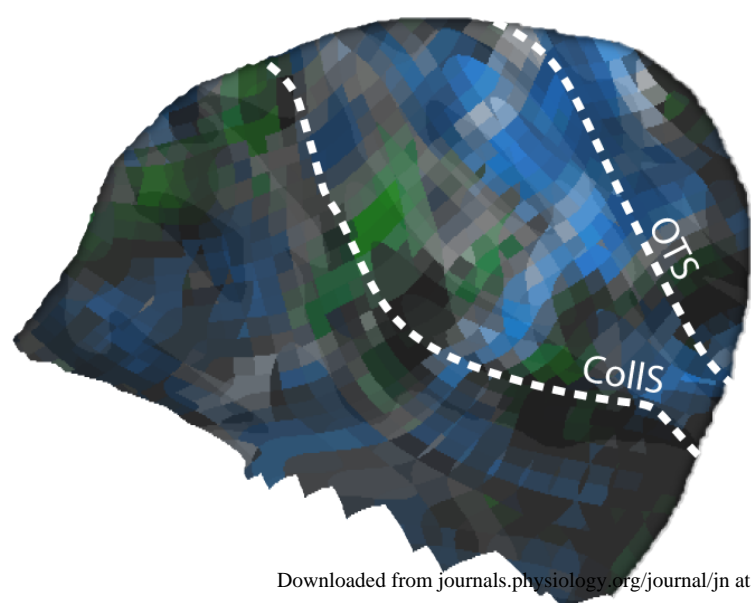727    indicate that brain representations are better explained by orthogonal components of variance.

34

**A**

| Curvature | Silhouette | Medial-axis | Inked area | Object Identity |

**B**

0 — Dissimilarity — 1
(n.u.)

fish rabbit shoe ball telephone teapot fish rabbit shoe ball fish rabbit shoe ball

teapot + telephone +

**C**

Curvature RDM

Silhouette RDM

Medial-axis RDM

Inked area RDM

Identity RDM

Searchlight

fMRI pattern RDM

Stimulus (1000ms)

Variance partitioning

$$R^2_{curvature} = R^2_{full} - R^2_{silhouette + medial\text{-}axis + inked\ area + object\ identity}$$

**D**

Medial-axis
Curvature
Object identity
Inked area

Silhouette  Medial-axis  Curvature  Object identity

-0.35 — Correlation — 0.35
(Pearson's r)

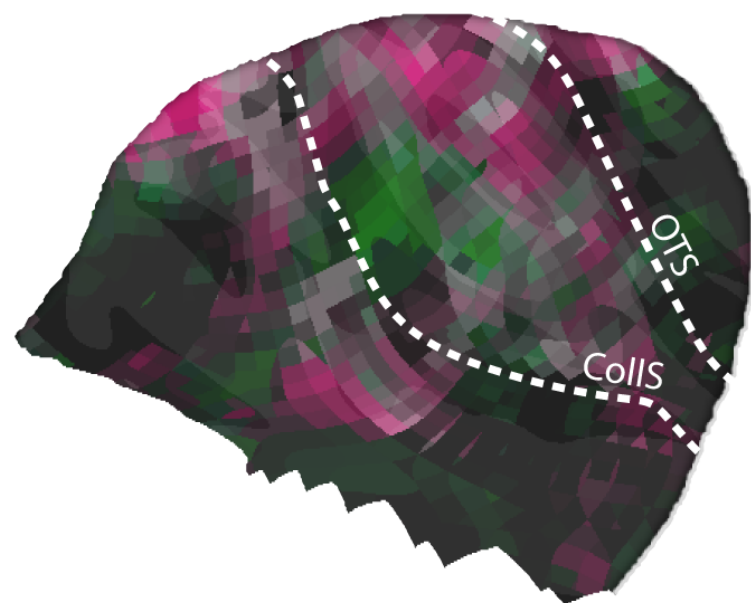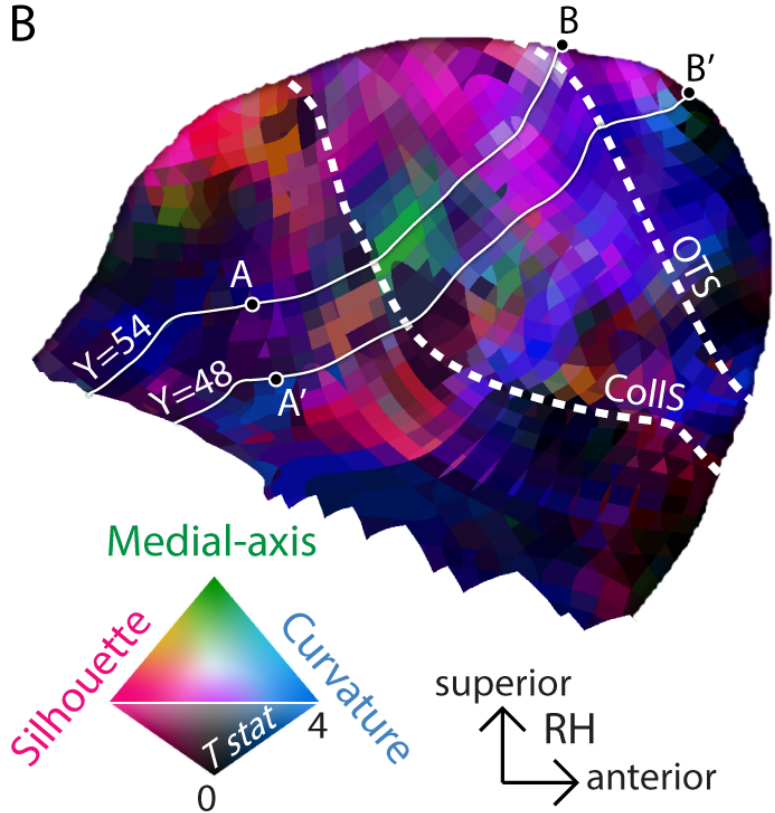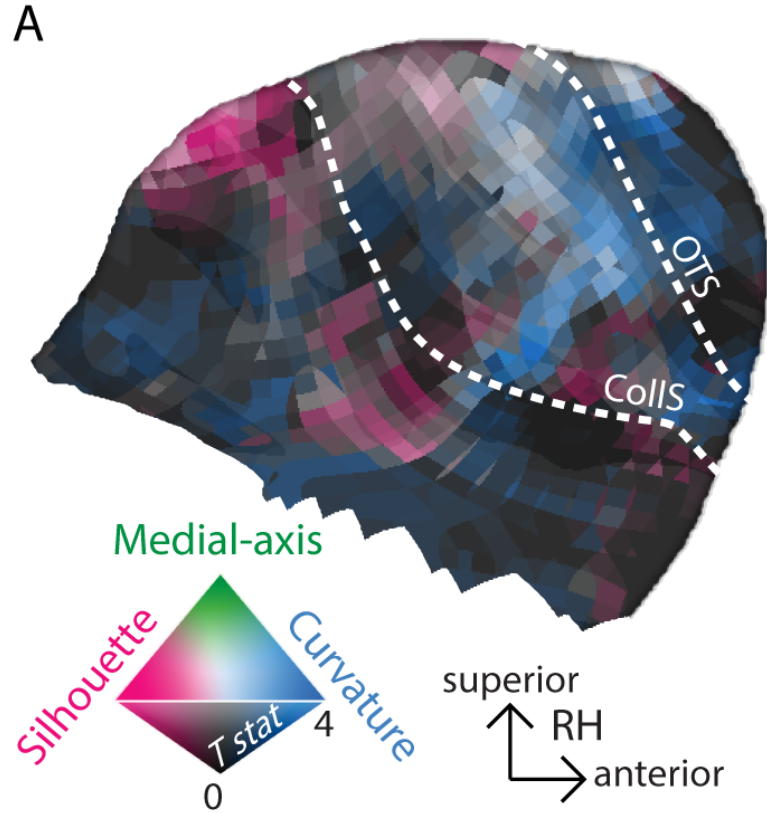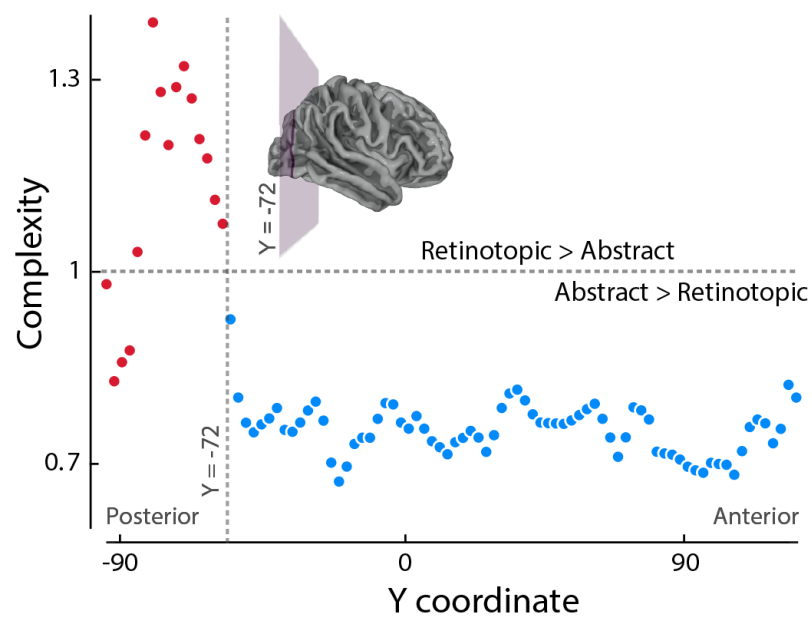Z = -23    Z = -13    Z = -3    Z = 7    Z= 17    Z = 27    Z = 37

L    R    ■ Curvature    ■ Silhouette    ■ Medial-axis    ■ Inked area    ■ Object Identity

Medial-axis

Silhouette

Curvature

T stat

0

4

IPS

STS

ITS

LO

OTS

CollS

IPS

STS

ITS

LO

OTS

CollS

IPS

STS

ITS

LO

OTS

CollS

IPS

STS

ITS

LO

OTS

CollS

IPS

STS

ITS

LO

OTS

CollS

IPS

STS

ITS

LO

OTS

CollS

superior

LH

anterior

superior

RH

anterior

**A**

Complexity

1.3

1

0.7

Retinotopic > Abstract

Abstract > Retinotopic

Y = -72

Posterior

Anterior

-90

0

90

Y coordinate

**B**

Orthogonality

2.3

2.15

$\rho = -0.83$ ***
Orthogonal > Shared

Posterior

Anterior

-90

0

90

Y coordinate