



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/153410/>

Version: Accepted Version

Article:

Kaiser, Daniel, Azzalini, Damiano C and Peelen, Marius V (2016) Shape-independent object category responses revealed by MEG and fMRI decoding. *Journal of Neurophysiology*. pp. 2246-50. ISSN: 0022-3077

<https://doi.org/10.1152/jn.01074.2015>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

1 **Shape-independent object category responses revealed by MEG and fMRI decoding**

2

3 Daniel Kaiser^{1*}, Damiano C. Azzalini^{1*}, Marius V. Peelen¹

4 *Center for Mind/Brain Sciences, University of Trento, 38068 Rovereto (TN), Italy*

5 **D.K. and D.C.A. contributed equally to this study*

6

7 **Corresponding Author**

8 Marius V. Peelen

9 Center for Mind/Brain Sciences, University of Trento

10 Corso Bettini 31, 38068 Rovereto (TN), Italy

11 Phone: +39 0464 808718, e-mail: marius.peelen@unitn.it

12

13 **Running Head**

14 Decoding shape-independent object category responses

15

16 **Author Contributions**

17 D.K., D.C.A., and M.V.P. conception and design of research; D.K. and D.C.A. performed
18 experiments; D.K. and D.C.A. analysed data; D.K., D.C.A., and M.V.P. interpreted results of
19 experiments; D.K. prepared figures; D.K. and M.V.P. drafted manuscript; D.K., D.C.A. and
20 M.V.P. edited and revised manuscript; D.K., D.C.A. and M.V.P. approved final version of
21 manuscript.

22 **Abstract**

23 Neuroimaging research has identified category-specific neural response patterns to a
24 limited set of object categories. For example, faces, bodies, and scenes evoke activity
25 patterns in visual cortex that are uniquely traceable in space and time. It is currently
26 debated whether these apparently categorical responses truly reflect selectivity for
27 categories or instead reflect selectivity for category-associated shape properties. In the
28 present study, we used a cross-classification approach on fMRI and MEG data to reveal
29 both category-independent shape responses and shape-independent category
30 responses. Participants viewed human body parts (hands and torsos) and pieces of
31 clothing that were closely shape-matched to the body parts (gloves and shirts).
32 Category-independent shape responses were revealed by training multivariate classifiers
33 on discriminating shape within one category (e.g., hands versus torsos) and testing these
34 classifiers on discriminating shape within the other category (e.g., gloves versus shirts).
35 This analysis revealed significant decoding in large clusters in visual cortex (fMRI),
36 starting from 90ms after stimulus onset (MEG). Shape-independent category responses
37 were revealed by training classifiers on discriminating object category (bodies, clothes)
38 within one shape (e.g., hands versus gloves) and testing these classifiers on
39 discriminating category within the other shape (e.g., torsos versus shirts). This analysis
40 revealed significant decoding in bilateral occipitotemporal cortex (fMRI), and from 130 to
41 200ms after stimulus onset (MEG). Together, these findings provide evidence for
42 concurrent shape and category selectivity in high-level visual cortex, including category-
43 level responses that are not fully explicable by 2D shape properties.

44 **Keywords:** Category Selectivity, Visual Cortex Organization, Body Representations

45 **Introduction**

46 Functional magnetic resonance imaging (fMRI) studies have shown that multi-voxel
47 response patterns in high-level visual cortex reliably discriminate different object
48 categories (Haxby et al., 2001), and that these show a meaningful categorical
49 organization (e.g., an animate-inanimate distinction; Kriegeskorte et al., 2008). Similarly,
50 signatures of category-specific processing in the time domain have been identified using
51 magneto- and electroencephalography (MEG/EEG), with MEG sensor patterns across the
52 scalp allowing for reliable classification of object categories (Carlson et al., 2013; Cichy et
53 al., 2014).

54 However, it is unclear whether such categorical responses are truly reflecting
55 category membership, detached from specific visual features, or whether they are
56 instead driven by visual properties of objects that systematically covary with category
57 membership. For example, the face-selective fusiform face area (Kanwisher et al., 1997)
58 is preferentially activated for round, non-face stimuli that have a higher spatial
59 concentration of elements in the upper half even when these stimuli are not recognized
60 as faces (Caldara et al., 2006), and the occipital face area (Gauthier et al., 2000) has been
61 shown to be causally involved in the perception of stimulus symmetry (Bona et al., 2015).
62 Furthermore, large-scale response patterns in monkey IT can be well explained by the
63 objects' shape similarity without the need to refer to category membership (Baldassi et
64 al., 2013). Such findings prompt the hypothesis that closely matching shape properties of
65 objects from different categories would largely abolish category-specific response
66 patterns.

67 We tested this prediction by investigating how matching for 2D shape properties
68 impacts neural responses to a specific category -- the human body. Previous studies have

69 characterized distinct spatio-temporal signatures of body perception, recruiting specific
70 regions in occipitotemporal and fusiform cortices and evoking specific
71 electrophysiological waveform components (for review, see Peelen and Downing, 2007).
72 Furthermore, bodies can be reliably separated from other categories based on MEG and
73 fMRI response patterns (Cichy et al., 2014; Kriegeskorte et al., 2008). It is unknown
74 whether these body-specific fMRI and MEG responses reflect selectivity for particular
75 shape properties of bodies (e.g., symmetry) or whether they reflect, at least partly, a
76 truly categorical response.

77 Participants were tested in separate fMRI and MEG experiments with largely
78 identical experimental procedures. Multivariate classification techniques were used to
79 characterize category representations in space (fMRI) and time (MEG). The stimulus set
80 consisted of human body parts (hands and torsos) and pieces of clothing (gloves and
81 shirts) that were closely shape-matched to the body part stimuli. To reveal category-
82 independent shape responses, classifiers were trained to discriminate between different
83 shapes within one category (e.g., hands versus torsos), and tested to discriminate these
84 shapes within the other category (e.g., gloves versus shirts). To reveal shape-
85 independent category responses, classifiers were trained to discriminate between the
86 categories (bodies, clothes) within one shape (e.g., hands versus gloves), and tested to
87 discriminate these categories within the other shape (e.g., torsos versus shirts).

88 **Materials & Methods**

89 *Participants.* Twenty-four healthy adults (11 male; mean age 24.2 years, SD=3.4)
90 took part in the fMRI experiment and 21 healthy adults (14 male; mean 25.0 years, SD=3.2)
91 took part in the MEG experiment. One participant completed both experiments. All
92 participants had normal or corrected-to-normal visual acuity. All procedures were carried
93 out in accordance with the Declaration of Helsinki and were approved by the ethical
94 committee of the University of Trento.

95 *Stimuli and Procedure.* Unless otherwise noted, all aspects of the design were
96 identical between the fMRI and MEG experiments. The full stimulus set consisted of nine
97 different categories (hands, gloves, torsos, shirts, brushes, pens, trees, vegetables, and
98 chairs), with 21 different exemplars per category. Four of these categories (brushes,
99 pens, trees, vegetables) were related to a different research question, and are not
100 analyzed here. Chairs served as target stimuli (see Fig. 1b), and were also excluded from
101 all analyses. Our analyses were focused on the comparison between stimuli depicting
102 human body parts (human hands and torsos, i.e. shirts with a human upper body inside)
103 and stimuli depicting solely pieces of clothing despite being very similar to the human
104 body parts in their shape properties (gloves and shirts; Fig. 1a).

105

106 >> Figure 1 <<

107

108 Both experiments consisted of multiple runs, where participants viewed grey-
109 scale images of the different categories for 500ms in a randomized order (see Fig. 1b),
110 with stimuli being separated by a fixation interval varying randomly between 1500 and
111 2000ms (in discrete steps of 50ms). Participants were instructed to maintain central

112 fixation and press the response button whenever they saw a chair (these trials appeared
113 equally often as all other categories, e.g., 21 times per run). For the MEG experiment,
114 participants were additionally instructed to specifically use the chair trials for eye blinks.
115 Each run contained each individual exemplar of every category exactly once, leading to a
116 total of 189 trials per run and an average run duration of 7.1 minutes. In the fMRI
117 experiment, every run additionally contained a 10s fixation period at the beginning and
118 end. During the fMRI experiment, participants completed six of these runs (for one
119 participant only data from five runs was collected due to a technical problem), and
120 during the MEG experiment, participants completed ten runs (one participant performed
121 eleven runs). Stimulus presentation was controlled using the Psychtoolbox (Brainard,
122 1997); in the MRI stimuli were back-projected onto a screen at the end of the scanner
123 bore and participants saw the stimulation through a tilted mirror mounted on the head
124 coil, while in the MEG, stimuli were back-projected onto a translucent screen located in
125 front of the participant.

126 *fMRI data acquisition and preprocessing.* MR imaging was conducted using a
127 Bruker BioSpin MedSpec 4T head scanner (Bruker BioSpin, Rheinstetten, Germany),
128 equipped with an eight-channel head coil. During the experimental runs, T2*-weighted
129 gradient-echo echo-planar images (EPI) were collected (repetition time TR=2.0s, echo
130 time TE=33ms, 73° flip angle, 3 x 3 x 3mm voxel size, 1mm gap, 34 slices, 192mm field of
131 view, 64x64 matrix size). Additionally, a T1-weighted image (MPRAGE; 1 x 1 x 1mm voxel
132 size) was obtained as a high-resolution anatomical reference. All resulting data were
133 preprocessed using MATLAB and SPM8. The functional volumes were realigned and
134 coregistered to the structural image. Additionally, structural images were spatially
135 normalized to the MNI-305 template (as included in SPM8), to obtain normalizing

136 parameters for each participant. These parameters were later used to normalize
137 individual participants' searchlight result maps before entering them into statistical
138 analysis.

139 *fMRI decoding analysis.* Multivariate pattern analysis (MVPA) was carried out on a
140 TR-based level using the CoSMoMVPA toolbox (www.cosmomvpa.org). To reveal areas
141 yielding above chance decoding throughout the brain, a searchlight analysis was
142 conducted, where a spherical neighborhood of 40 voxels (6.4mm average radius) was
143 moved across the whole brain. For each voxel belonging to a specific neighborhood, TRs
144 corresponding to the conditions of interested were selected by shifting the voxel-wise
145 time-course of activation by three TRs (to account for the hemodynamic delay).
146 Subsequently, for each run separately, activation values were extracted from the
147 unsmoothed EPI-volumes for each TR coinciding with the onset of a specific condition.
148 Similar to the MEG analysis, MVPA was done in a pairwise fashion: Linear discriminant
149 analysis (LDA) classifiers were trained to discriminate response patterns for two
150 conditions in all but one runs and were subsequently tested on response patterns for
151 these two conditions taken from the remaining, left-out run. This procedure was
152 repeated, so that every run served as the testing set once. For the cross-decoding
153 analysis, classifiers were trained on discriminating two conditions (e.g. hands versus
154 gloves), and tested on two different conditions (e.g. torsos versus shirts); thus, for this
155 analysis, all available trials were used in the training and test set. Pairwise classification
156 accuracy for every voxel was assessed by comparing the labels predicted by the classifier
157 to the actual labels, with chance performance always being 50%. Individual-subject
158 searchlight maps were normalized to MNI-space before they were entered into statistical
159 analyses. Above-chance classification was identified using a threshold-free cluster

160 enhancement (TFCE) procedure (Smith and Nichols, 2009), where the observed decoding
161 accuracy was tested against a simulated null-distribution (generated from 10,000
162 bootstrapping iterations). The resulting statistical maps were thresholded at $p < 0.05$
163 (one-tailed).

164 *MEG acquisition and preprocessing.* Electromagnetic brain activity was recorded
165 using an Elekta Neuromag 306 MEG system (Elekta Neuromag® systems, Helsinki,
166 Finland), composed of 204 planar gradiometers and 102 magnetometers. Signals were
167 sampled continuously at 1000 Hz and band-pass filtered online between 0.1 and 330Hz.
168 Offline preprocessing was done using MATLAB and the fieldtrip analysis package
169 (Oostenveld et al., 2011). Data were concatenated for all runs, high-pass filtered at 1Hz,
170 and epoched into trials ranging from -100 to 500ms with respect to stimulus onset. Based
171 on visual inspection, trials containing eye blinks and other movement-related artifacts
172 were completely discarded from all analyses. Data was then baseline-corrected with
173 respect to the pre-stimulus window and downsampled to 100 Hz to increase the signal-
174 to-noise ratio of the multivariate classification analysis (see Carlson et al., 2013).

175 *MEG decoding analysis.* MVPA was carried out on single trial data using the
176 CoSMoMVPA toolbox (www.cosmomvpa.org). Only magnetometers were used, as these
177 sensors allowed for the most reliable classification in previous work in our lab (Kaiser et
178 al., 2015). Classification was performed using LDA classifiers. For the shape cross-
179 decoding analysis, classifiers were trained on one category-matched shape comparison
180 (i.e., hands versus torsos or gloves versus shirts) and tested on the other comparison
181 (i.e., gloves versus shirts or hands versus torsos). For the category cross-decoding
182 analysis, classifiers were trained on one shape-matched category comparison (i.e., hands
183 versus gloves or torsos versus shirts) and tested on the other comparison (i.e., torsos

184 versus shirts or hands versus gloves). To increase the reliability of the data supplied to
185 the classifiers, new, “synthetic” trial data was created by averaging single trial data
186 separately for every condition and chunk, by randomly picking 25% of trials and averaging
187 this data across trials. This procedure was repeated 100 times (with the constraint that
188 no trial was used more than one time more often than any other trial), so that for every
189 condition and chunk, 100 of these synthetic trials were available for classification.
190 Classification accuracy was then assessed by computing the percentage of correctly
191 classified trials in the test chunk, with chance performance being 50%. Classification was
192 repeated for every possible combination of training and testing time points, leading to a
193 60 X 60 time points (600 X 600ms, with 100Hz temporal resolution) matrix of
194 classification accuracies. Individual subject accuracy maps were smoothed using a 3 X 3
195 time points (i.e. 30ms in train and test time) averaging filter. To identify time periods of
196 significant above-chance classification, similar to the fMRI analysis, a TFCE procedure was
197 used, where the observed decoding accuracy was tested against a simulated null-
198 distribution (generated from 10,000 bootstrapping iterations). The baseline (pre-
199 stimulus) interval was not taken into account for statistical testing. The resulting
200 statistical maps were thresholded at $p < 0.05$ (one-tailed).
201

202 **Results**

203 *Shape Cross-Decoding.* Brain regions representing object shape across categories
204 were identified by training classifiers on discriminating shape within one category (e.g.,
205 hands versus torsos), and testing these classifiers on discriminating shape within the
206 other category (e.g., gloves versus shirts). Results from both possible train/test-
207 directions were averaged. An fMRI searchlight using this approach revealed regions in
208 right (33,128mm³; peak MNI coordinate: x=48, y=-68, z=-4; $t_{23}=8.5$) and left (30,368mm³;
209 peak MNI coordinate: x=-6, y=-94, z=-12; $t_{23}=9.6$) visual cortex, spanning early visual areas
210 and regions of lateral occipitotemporal cortex (Fig. 2a,b). The MEG data showed above-
211 chance decoding of shape, starting at 90ms after stimulus onset, and peaking along the
212 diagonal at 170ms and 240ms (467 time points in total, maximum decoding accuracy:
213 70.2%; $t_{20}=11.4$; Fig. 2c).

214

215 >> Figure 2 <<

216

217 *Category Cross-Decoding.* A second cross-decoding analysis was conducted to test
218 for responses that reflect object category (body parts versus clothes), independently of
219 shape properties. To detect such shape-independent responses, classifiers were trained
220 to discriminate bodies and clothes for one shape-matched comparison (e.g., hand versus
221 glove), and subsequently tested on the other comparison (e.g., torso versus shirt).
222 Results from both possible train/test-directions were averaged.

223

224 >> Figure 3 <<

225

226 In the fMRI searchlight analysis, clusters in right (3,664mm³; peak MNI coordinate:
227 x=52, y=-70, z=6; t_{23} =6.5) and left (5,752mm³; peak MNI coordinate: x=-44, y=-78, z=10;
228 t_{23} =5.8) lateral occipitotemporal cortex were identified (Fig. 3a). These clusters
229 overlapped with the extrastriate body area (EBA; Fig. 3b; coordinates of Downing et al.,
230 2001: x=±/51, y=-72, z=5). Performing the same cross-classification analysis on the MEG
231 data revealed a specific temporal signature associated with shape-independent category
232 responses: classifiers could reliably discriminate between bodies and clothes between
233 130–160ms with respect to the hand-glove comparison and 160–200ms with respect to
234 the torso-shirt comparison (12 time points in total, maximum decoding accuracy: 53.6%;
235 t_{20} =6.9; Fig. 3c).

236 **Discussion**

237 Here we asked whether categorical representations in visual cortex are fully driven by
238 category-associated visual features or if they (at least partly) reflect category
239 membership. Unlike previous studies investigating category selectivity, the stimuli
240 presented in the current study were matched for shape properties, including object-part
241 structure (e.g. hands and gloves both have five "fingers"), outline similarity, and
242 symmetry. We found that large clusters in visual cortex are sensitive to shape differences
243 (i.e., "hand/glove"-shape versus "upper body"-shape): classifiers trained on
244 discriminating hands and torsos successfully discriminated gloves and shirts (and vice
245 versa), in both early visual areas and occipitotemporal cortex. These shape differences
246 were reliably decodable from MEG response patterns as early as 90ms after stimulus
247 onset.

248 Crucially, we also found evidence for shape-independent category responses:
249 classifiers trained on discriminating hands and gloves successfully discriminated torsos
250 and shirts (and vice versa) in bilateral clusters in the occipitotemporal cortex. These large
251 clusters likely encompass body-, motion-, and object-selective regions of visual cortex,
252 which closely overlap both at the group-level and within individual subjects (Downing et
253 al., 2007). Interestingly, the MEG data showed a specific temporal profile associated with
254 such shape-independent body representations. Response patterns between 130 and
255 200ms after stimulus onset allowed for successful cross-classification, in line with
256 previous electrophysiological findings showing that bodies can be differentiated from
257 other categories based on scalp distributions from 130 to 230 ms (Thierry et al., 2006).
258 These fMRI and MEG results thus confirm previous studies on body-selective responses
259 but additionally show that this selectivity is not fully explicable by 2D shape properties.

260 A particular strength of the cross-decoding approach used here is that it provides
261 a rigorous control of possible visual differences between the two categories (bodies,
262 clothes), beyond the shape matching of the two body-clothing pairs: uncontrolled visual
263 differences in one comparison (e.g., the presence of a neck in torsos, not shirts) would
264 also need to be present in the other comparison (e.g., hand versus glove) for these
265 differences to lead to successful decoding. Thus, successful decoding in this analysis
266 likely reflects genuine category membership rather than visual or shape properties.
267 Similarly, it is unlikely that differences in the deployment of spatial attention could
268 account for the results: classifiers picking up on such differences between the two
269 training stimuli (e.g., a preferential allocation of attention to the upper part of torsos, but
270 not shirts) are unlikely to benefit from this when tested on the other comparison. It is still
271 possible, in principle, that there are remaining visual differences, such as skin texture or
272 3D volume, that are shared by the body conditions but not the clothes conditions.
273 However, we think it is unlikely that such features would drive body-selective responses,
274 considering previous work showing body-selective responses to highly schematic
275 depictions of the body lacking these cues (e.g., point-light motion, stick figures,
276 silhouettes; Peelen and Downing, 2007). Nevertheless, further studies are needed to
277 identify and rule out any such remaining differences.

278 We interpret the present findings as showing that the presence of particular
279 visual or shape features is not necessary for evoking a body-selective response. Rather,
280 these responses appear to reflect (or follow from) the categorization of an object as
281 being a body part – a category that is associated with specific perceptual and conceptual
282 properties, such as bodily actions/movements, social relevance, and agency (Sha et al.,
283 2014). Different cues can support the inference that a perceived object is a body. These

284 cues are often part of the object itself (e.g., characteristic body shapes or movements)
285 but may also come from the surrounding context (Cox et al., 2004), from other
286 modalities, or from expectations and knowledge (e.g., knowing that a mannequin in a
287 shopping window is not a human). Our results show that body-selective responses in
288 lateral occipitotemporal cortex, emerging at around 130-200ms, follow from this
289 categorical inference rather than reflecting a purely stimulus-driven response to the
290 visual features of the object.

291 Interestingly, clusters exhibiting category-independent shape responses
292 overlapped with clusters exhibiting shape-independent category responses. This
293 observation is congruent with previous studies highlighting both visual (Andrews et al.,
294 2015; Baldassi et al., 2013) and semantic (Huth et al., 2012; Sha et al., 2014) dimensions as
295 organizational principles of high-level visual cortex. Response patterns in inferotemporal
296 cortex seem to be best explicable by models using a combination of visual feature
297 attributes and category membership (Khaligh-Razavi and Kriegeskorte, 2014), suggesting
298 that in high-level visual cortex these representations co-exist.

299 While the fMRI data demonstrated that shape and category responses are
300 spatially entwined, the MEG results revealed differing temporal dynamics of these
301 responses: while shape-specific responses could be decoded early and across a relatively
302 long time interval, shape-independent category responses showed a specific temporal
303 signature between 130 and 200ms. We interpret this as a temporally restricted period
304 where cortical responses reflect processing of category membership: successful
305 decoding in the category cross-decoding analysis does not only require shape-
306 independence of body-specific responses, but also generalization across different body
307 parts. This generalization might be restricted to the specific time window revealed here,

308 with earlier computations reflecting stimulus-specific attributes (related to individual
309 body parts), and later processing reflecting more sophisticated stimulus analysis that
310 diverges for different body parts (e.g., hands carry different social and action-related
311 information than torsos). Hence, the temporally specific generalization across body parts
312 observed here might reflect a unique timestamp of category-level recognition.
313 Interestingly, this category-level recognition occurred at different time points for the two
314 body parts included in the study, with slightly faster categorization of the hands (130-
315 160ms) than the torso (160-200ms). This later discriminability of torsos and shirts may
316 reflect the greater similarity of these two stimuli on a perceptual level (see Fig. 1a),
317 leading to relatively delayed recognition of the torsos as being a body part.

318 To conclude, the present study characterizes the spatial and temporal profiles of
319 shape-independent categorical neural responses by showing that MEG and fMRI
320 response patterns distinguish between body parts and closely matched control stimuli.
321 The patterns that distinguished each of the two body parts from their respective shape-
322 matched controls showed sufficient commonality to allow for cross-pair decoding of
323 object category. These generalizable category-selective response patterns were localized
324 in space (lateral occipitotemporal cortex) and time (130-200ms after stimulus onset).

325 **Acknowledgements**

326 We thank Nick Oosterhof for help with data analysis.

327

328 **Grants**

329 The research was funded by the Autonomous Province of Trento, Call "Grandi Progetti

330 2012", project "Characterizing and improving brain mechanisms of attention - ATTEND".

331 **References**

- 332 **Andrews TJ, Watson DM, Rice GE, Hartley T.** Low-level properties of natural images
333 predict topographic patterns of neural responses in the ventral visual pathway. *J*
334 *Vis*, 15: 3, 2015.
- 335 **Baldassi C, Alemi-Neissi A, Pagan M, Dicarlo JJ, Zecchina R, Zoccolan D.** Shape similarity,
336 better than semantic membership, accounts for the structure of visual object
337 representations in a population of monkey inferotemporal neurons. *PLoS Comput*
338 *Biol*, 9: e1003167, 2013.
- 339 **Bona S, Cattaneo Z, Silvanto J.** The causal role of the occipital face area (OFA) and lateral
340 occipital (LO) cortex in symmetry perception. *J Neurosci*, 35: 731-738, 2015.
- 341 **Brainard DH.** The psychophysics toolbox. *Spat Vis*, 10: 433-436, 1997.
- 342 **Caldara R, Seghier ML, Rossion B, Lazeyras F, Michel C, Hauert CA.** The fusiform face
343 area is tuned for curvilinear patterns with more high-contrasted elements in the
344 upper part. *Neuroimage*, 31: 313-319, 2006.
- 345 **Carlson T, Tovar DA, Alink A, Kriegeskorte N.** Representational dynamics of object vision:
346 The first 1000 ms. *J Vis*, 13: 1-19, 2013.
- 347 **Cichy RM, Pantazis D, Oliva A.** Resolving human object recognition in space and time. *Nat*
348 *Neurosci*, 17: 455-462, 2014.
- 349 **Cox D, Meyers E, Sinha P.** Contextually evoked object-specific responses in human visual
350 cortex. *Science*, 304: 115-117, 2004.
- 351 **Downing PE, Wiggett AJ, Peelen MV.** Functional magnetic resonance imaging
352 investigation of overlapping lateral occipitotemporal activations using multi-voxel
353 pattern analysis. *J Neurosci*, 27: 226-233, 2007.
- 354 **Downing PE, Jiang Y, Shuman M, Kanwisher N.** A cortical area selective for visual
355 processing of the human body. *Science*, 293: 2470-2473, 2001.

Formatted: German

356 **Gauthier I, Tarr MJ, Movlan J, Skudlarski P, Gore JC, Anderson AW.** The fusiform “face
357 area” is part of a network that processes faces at the individual level. *J Cogn*
358 *Neurosci*, 12: 495-504, 2000.

359 **Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P.** Distributed and
360 overlapping representations of faces and objects in ventral temporal cortex.
361 *Science*, 293: 2425-2430, 2001.

362 **Huth AG, Nishimoto S, Vu AT, Gallant JL.** A continuous semantic space describes the
363 representation of thousands of object and action categories across the human
364 brain. *Neuron*, 76: 1210-1224, 2012.

365 **Julian JB, Fedorenko E, Webster J, Kanwisher N.** An algorithmic method for functionally
366 defining regions of interest in the ventral visual pathway. *Neuroimage*, 60: 2357-
367 2364, 2012.

368 **Kanwisher N, McDermott J, Chun MM.** The fusiform face area: a module in human
369 extrastriate cortex specialized for the perception of faces. *J Neurosci*, 17: 4302-
370 4311, 1997.

371 **Kaiser D, Oosterhof NN, Peelen MV.** The temporal dynamics of target selection in real-
372 world scenes. *J Vis*, 15: 740, 2015.

373 **Khaligh-Razavi S-M, Kriegeskorte N.** Deep supervised, but not unsupervised, models may
374 explain IT cortical representation. *PLoS Comput Biol*, 10: e1003915, 2014.

375 **Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini P.**
376 Matching categorical object representations in inferior temporal cortex of man
377 and monkey. *Neuron*, 60: 1126-1141, 2008.

Formatted: English (UK)

378 **Oostenveld R, Fries P, Maris E, Schoffelen J-M.** FieldTrip: Open source software for
379 advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput*
380 *Intell Neurosci*, 2011: 156869, 2011.

381 **Peelen MV, Downing PE.** The neural basis of visual body perception. *Nat Rev Neurosci*, 8:
382 636-648, 2007.

383 **Sha L, Haxby JV, Abdi H, Guntupalli JS, Oosterhof NN, Halchenko YO, Connolly AC.** The
384 animacy continuum in the human ventral vision pathway. *J Cogn Neurosci*, 37: 665-
385 678, 2015.

386 **Smith SM, Nichols TE.** Threshold-free cluster enhancement: Addressing problems of
387 smoothing, threshold dependence and localisation in cluster inference.
388 *Neuroimage*, 44: 83-98, 2009.

389 **Thierry G, Pegna AJ, Dodds C, Roberts M, Basan S, Downing PE.** An event-related
390 potential component sensitive to images of the human body. *Neuroimage*, 32: 871-
391 879, 2006.

392 **Figure Legends**

393

394 Fig. 1. Stimuli and Paradigm. a) The stimulus set contained two human body parts (hands
395 and torsos; see first and third row for examples), and two pieces of clothing that are
396 highly similar in their shape (gloves and shirts; second and fourth row). b) Stimuli were
397 presented for 500ms, separated by a variable 1500 – 2000ms fixation interval.

398 Participants were instructed to maintain central fixation and to manually respond to
399 chairs.

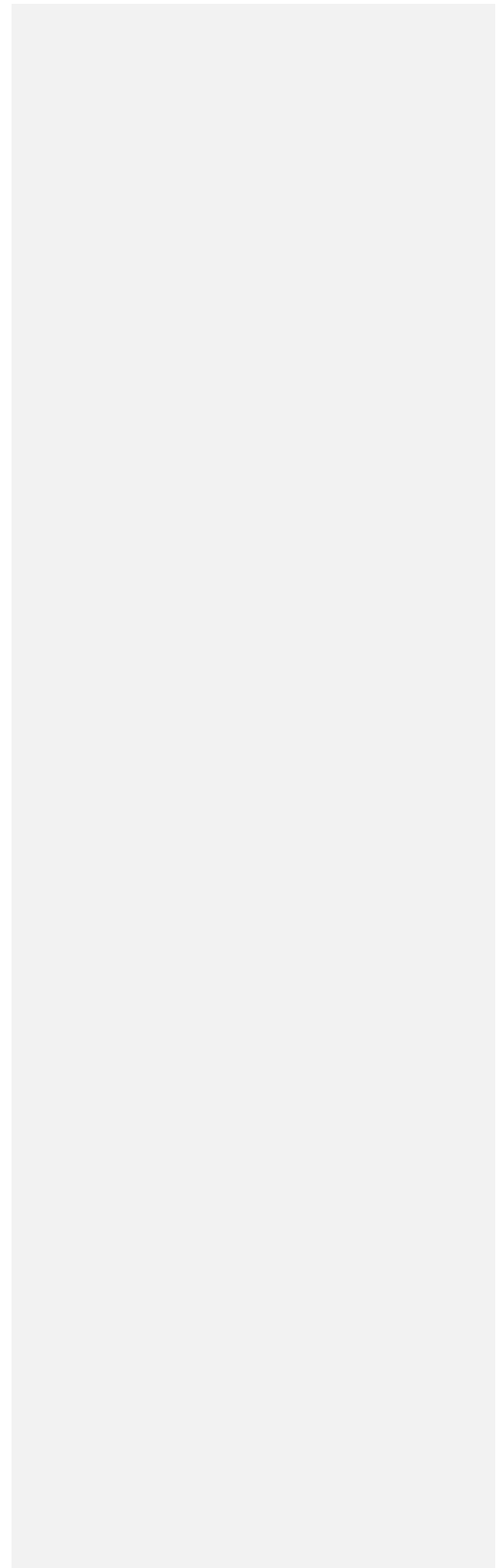
400

401 Fig. 2. Shape cross-decoding analysis. To reveal shape-selective mechanisms, classifiers
402 were trained to discriminate shape within one category (e.g. hands versus torsos), and
403 tested on the other category (e.g. gloves versus shirts). Results from both train/test
404 directions were averaged. a,b) fMRI decoding was significantly above chance in large
405 areas of visual cortex, spanning primary visual areas and regions of occipitotemporal
406 cortex. c) MEG decoding was significantly above chance along the diagonal, starting from
407 90ms after stimulus onset and peaking after 170ms and 240ms. Note that the axes here
408 reflect time with respect to the two possible train and test comparisons, independently
409 of the actual train/test-direction. The connected area indicates above-chance decoding.

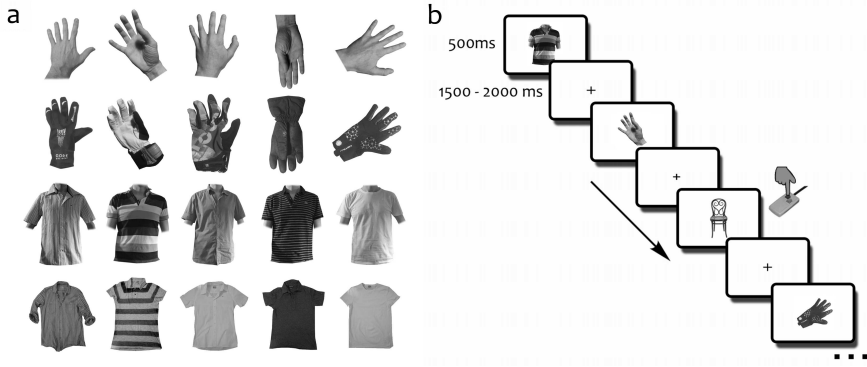
410

411 Fig. 3. Category cross-decoding analysis. To reveal generalization across the two body-
412 clothes pairs, classifiers were trained on one comparison (e.g. hands versus gloves), and
413 tested on the other (e.g. hands versus gloves). Results from both train/test directions
414 were averaged. a) fMRI decoding was significantly above chance in bilateral regions of
415 lateral occipito-temporal cortex. b) The clusters obtained in this searchlight analysis fell

416 within regions previously reported as body-selective -- the black outline represents the
417 boundaries of a group map of body-selectivity in occipitotemporal cortex (taken from
418 <http://web.mit.edu/bcs/nklab/GSS.shtml>). c) MEG decoding revealed a temporally specific
419 window of successful cross-classification ranging from 130–160ms with respect to the
420 hand-glove comparison (“hand time”) and from 160–200ms with respect to the torso-
421 shirt comparison (“torso time”). Note that the axes here reflect time with respect to the
422 two possible train and test comparisons, independently of the actual train/test-direction.
423 The connected area indicates above-chance decoding.



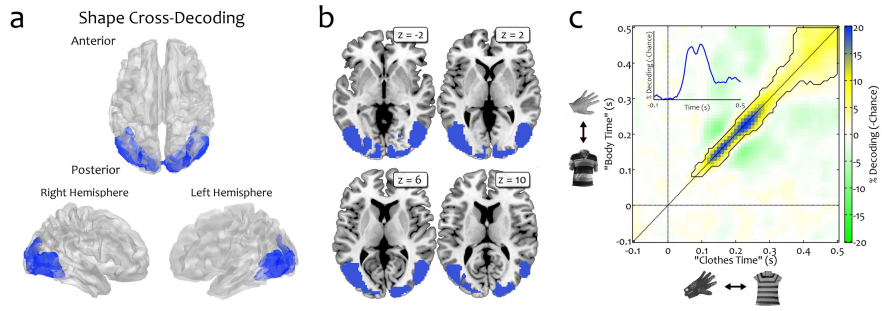
424 **Figure 1**



425

426

427 **Figure 2**



428

429

430 **Figure 3**

