

1 **Title: The representation of colored objects in macaque color patches**

2 **Authors: Le Chang¹, Pinglei Bao¹ and Doris Y. Tsao^{1,2}**

3 **Affiliations:**

4 1. Division of Biology and Biological Engineering, Computation and Neural Systems, Caltech,
5 Pasadena CA 91125.

6 2. Howard Hughes Medical Institute, Pasadena, CA 91125.

7 **Correspondence:**

8 Doris Y. Tsao: dortsao@caltech.edu

9

10 **Abstract:**

11 An important question about color vision is: how does the brain represent the color of an object? The
12 recent discovery of “color patches” in macaque inferotemporal (IT) cortex, the part of brain responsible
13 for object recognition, makes this problem experimentally tractable. Here we record neurons in three
14 color patches, middle color patch CLC (central lateral color patch), and two anterior color patches ALC
15 (anterior lateral color patch) and AMC (anterior medial color patch), while presenting images of objects
16 systematically varied in hue. We found that all three patches contain high concentrations of hue-
17 selective cells, and the three patches use distinct computational strategies to represent colored objects:
18 while all three patches multiplex hue and shape information, shape-invariant hue information is much
19 stronger in anterior color patches ALC/AMC than CLC; furthermore, hue and object shape specifically for
20 primate faces/bodies are over-represented in AMC but not in the other two patches.

21

22

23

24

25

26

27

28

29

30 Introduction

31 We see the world in color because different objects are composed of materials with different
32 reflectance spectra. The perception of color involves processing at multiple stages of the visual system.
33 Recordings in early parts of the visual system reveal double-opponent cells in area V1 ^{1,2} and hue-
34 selective cells in areas V2 and V4 ^{3,4}; these cells allow the brain to compute the local hue at each location
35 across a surface. But color processing does not end with extraction of local hue: the brain needs to
36 integrate information about hue distributions across space with information about large-scale object
37 shapes, to enable an organism to recognize and respond appropriately to colored objects. For example,
38 in Fig. 1a, we readily perceive a red apple, which requires (1) correctly discriminating the object from the
39 background, and (2) extracting the dominant hue within the object (Fig. 1a).

40 In theory, there are two possible mechanisms by which the brain could effectively organize information
41 about the color and shape of colored objects (we use the term “shape” to refer to all aspects of an
42 object’s identity independent of hue, not simply overall shape). Color and shape information could be
43 segregated into parallel channels, resulting in shape-invariant color-selective units and color-invariant
44 shape-selective units (Fig. 1b, top). Alternatively, units could be sharply tuned to both color and shape
45 (Fig. 1b, bottom). Our ability to name colors independent of shape (e.g., a “red” traffic light or a “red”
46 apple) supports the first scheme, while our ability to respond to specific color-shape combinations (e.g.,
47 stop at a red traffic light, eat a red apple) supports the second scheme (though it’s also possible such
48 semantic representations are not directly related to visual representations). In addition to asking how
49 single cells encode shape and color at different stages of visual processing, we can also ask how the
50 representation of colored objects at different stages of visual processing is transformed at the
51 population level: are shape and color co-represented within intermingled cell populations along the
52 entire visual pathway, or do they become segregated into separate populations?

53 To clarify how the representation of object color is transformed in the visual system following the
54 extraction of local hue information at both the single cell and population level, we targeted fMRI-
55 identified “color patches” in inferotemporal (IT) cortex ⁵ for electrophysiological recordings in three
56 macaque monkeys. IT cortex has long been believed to be responsible for the representation of object
57 shape ^{6,7}, but a recent fMRI study revealed a set of regions in macaque IT selective for colored compared
58 to grayscale gratings ⁵. Interestingly, color patches were yoked in position to face patches ⁵, regions in IT
59 cortex selective for faces ⁸. The localization of color patches within IT cortex implicates their role in
60 representing color within the context of object recognition, while the stereotyped relative localization of
61 color patches and face patches raises the possibility that a hierarchical functional organization for
62 processing colored objects exists, mirroring that for processing facial identity in the face patches ⁹.
63 Previous studies of color processing in monkeys have mostly used artificial stimuli, such as white noise,
64 sinusoidal gratings or simple geometric shapes ^{5,10-12}, while previous studies of object representation in

65 IT have mostly used grayscale images, ignoring color variations^{13,14}. This is unsatisfying, given that
66 *colored objects* are the only natural visual inputs. Here, we explored the co-representation of color and
67 object identity by presenting images of objects systematically varied in color, while recording from color
68 patches in IT.

69 **Results**

70 **Recording sites, connectivity and stimulus generation**

71 We first localized color patches in three monkeys with fMRI using colored versus black-white gratings⁵
72 (Fig. 1c). This revealed the central lateral, anterior lateral, and anterior medial color patches (CLC, ALC,
73 and AMC) in monkey M1, CLC, ALC, and the anterior fundus color patch (AFC) in monkey M2 (we could
74 not find AMC in this animal), and CLC, ALC, and AFC in monkey M3 (we could not find AMC in this
75 animal). Next, we electrically microstimulated color patches while performing simultaneous fMRI (see
76 Methods), to reveal the anatomical connectivity of color patches, and to potentially identify AMC in
77 monkeys M2 and M3. This technique has previously been used to study the connectivity of face patches
78¹⁵, and to reveal a place-selective region downstream of another place-selective region that had been
79 identified using fMRI¹⁶. Stimulating ALC in monkey M2's left hemisphere activated three additional
80 patches. Two of these patches overlapped with AFC and CLC identified by the color localizer (Fig. 1d),
81 while one of these patches was located anterior to the stimulation site, on the ventral surface of the
82 inferotemporal gyrus medial to the anterior middle temporal sulcus. Based on its location and
83 connectivity to ALC, we designated this patch AMC (Fig. 1c, d). Possibly this patch was actually missing
84 and we found another one, however, importantly, physiology was consistent between monkeys for AMC
85 identified in the two ways. Furthermore, stimulating CLC in the right hemisphere of monkey M2
86 activated a patch overlapping with ALC as identified by the color localizer (Supplementary Fig. 1). Overall,
87 these results suggest that color patches in IT cortex form a strongly and specifically interconnected
88 network, similar to face patches¹⁵, and allowed us to identify a color patch anterior to ALC in an animal
89 in which it was missing based on the color localizer experiments.

90 We next targeted middle color patch CLC, and anterior color patches ALC and AMC for
91 electrophysiological recordings. For comparison, we also targeted several sites in IT outside the color
92 patches, including face patch AM. To study the co-representation of color and object shape, we
93 generated a stimulus set of 82 images from 10 different categories (Supplementary Fig. 2a), each
94 rendered in 8 different hues (Fig. 1e; see Methods). In this way, we varied hue information and object
95 shape information independently and simultaneously. Grayscale images and the original color images
96 were also included in the stimulus set.

97 **Representation of hue by neurons in color patches**

98 Color stimuli were presented for 200 ms (ON period) interleaved by a gray screen for 200 ms (OFF
99 period) during recording. The full stimulus set was presented 7-10 times for each cell recorded.
100 Responses of all cells in each patch as well as cells outside the color patches to grayscale images and a
101 subset of color images are shown in Fig. 2. ANOVA analysis was performed to test the significance of
102 color tuning (see Methods). 74.7% (65/87) of cells recorded in CLC, 84.6% (55/65) of cells in ALC and
103 88.9% (80/90) of cells in AMC were significantly color tuned (ANOVA, $p < 0.001$); outside the color
104 patches, only 27.3% (9/33) of cells were significantly color tuned. Within color patches, only significantly
105 tuned cells were used for further analysis.

106 Responses to the 10 color conditions for all color-selective cells, grouped according to object category,
107 are shown in Fig. 3a. Each row represents one cell; cells in each patch are sorted according to hue
108 preference, computed using responses of each cell to 8 hues averaged across all 82 objects. There is
109 clear consistency of hue tuning across categories, especially for ALC and AMC. Hue consistency was
110 quantified by computing the correlation of hue tuning across categories (Supplementary Fig. 3a; see
111 Supplementary Fig. 3b for correlations between all pairs of categories). We found all neurons
112 demonstrated a positive correlation, with ALC and AMC significantly more consistent than CLC
113 ($W(65,55)=3215$, $p=2*10^{-4}$ between CLC and ALC; $W(65,80)=3473$, $p=4*10^{-7}$ between CLC and AMC;
114 $W(55,80)=3767$, $p=0.91$ between ALC and AMC, Wilcoxon rank sum test; average correlation value:
115 0.533 for CLC, 0.686 for ALC, 0.706 for AMC, and 0.119 for outside color patches). Outside the color
116 patches, consistency of hue tuning was much lower (Supplementary Fig. 3a, $W(33,200)=803$, $p=5*10^{-17}$,
117 Wilcoxon rank sum test).

118 We observed a difference between hue tuning in CLC/ALC compared to AMC. Most AMC cells preferred
119 red or yellow to other hues, leading to an under-representation of green and blue (Fig. 3a, arrow); in CLC
120 and ALC, different hues were more evenly represented (Fig. 3a; the proportion of cells preferring the 4
121 hues, from purple to green in CLC, ALC and AMC is 33.9 %, 47.3 % and 8.7 %. Chi-square test, $\chi^2(1)=12.61$
122 and $p=3*10^{-4}$ between CLC and AMC; $\chi^2(1)=23.11$ and $p=2*10^{-6}$ between ALC and AMC; $\chi^2(1)=2.22$ and
123 $p=0.13$ between CLC and ALC). Multidimensional scaling (MDS) analyses on population responses
124 revealed an additional difference in hue representation between AMC and the other two patches: In CLC
125 and ALC, the neural representation of all 8 hues is homogeneous, with gray located in the center of eight
126 hues; in AMC, yellow is over-represented, with gray located in the periphery, very close to cyan and
127 green (Fig. 3b). Note that gray is surrounded by the eight hues in the original CIE color space (Fig. 1e),
128 analogous to the population representation in ALC visualized with MDS.

129 We quantified the transformation in hue representation across CLC, ALC, and AMC by computing the
130 “neural” distance between neighboring hues based on population responses. AMC neurons displayed
131 stronger inhomogeneity in distance between neighbors than the other two patches (Fig. 3c). The
132 difference in color tuning between CLC/ALC and AMC was further confirmed by representation similarity

133 matrices quantifying the correlation between mean population responses to pairs of colors in each of
134 the three patches (Fig. 3d). The correlation between gray and cyan/green is evident in AMC, but almost
135 absent in the other two patches (mean correlation= 0.02 ± 0.08 for CLC; -0.02 ± 0.12 for ALC; 0.45 ± 0.09 for
136 AMC; $p<0.001$ between CLC and AMC; $p<0.001$ between ALC and AMC; $p=0.385$ between CLC and ALC,
137 20000 iterations of bootstrapping). In all three patches, the representation of the original color images
138 was closest to yellow/red (Fig. 3b, d). This could be due to the fact that pixel intensities of the original
139 color images turned out to be tightly distributed around yellow, especially for the faces and bodies
140 (Supplementary Fig. 2b-h). This last fact could also explain why AMC preferred red/yellow hues: it could
141 be biased to represent the color of faces and bodies. However, it's worth noting that AMC neurons
142 showed preference for red/yellow even for objects that are not naturally red/yellow, e.g. grapes,
143 watermelons, and abstract shapes without any natural color association (Fig. 3e). Thus AMC cells were
144 not simply over-representing correctly-colored objects.

145 One concern is that the preference for red/yellow observed in AMC could be due to undersampling in
146 our single-unit recordings. We recorded from AMC in ten different penetrations in two monkeys, and
147 results were consistent across both animals. To further address this concern, we performed an fMRI
148 experiment in which we presented red, yellow, blue, and grayscale monkey faces (see below for
149 rationale for showing monkey faces). Contrasting red/yellow versus grayscale monkey faces revealed
150 activation in CLC, ALC, and AMC (Supplementary Fig. 4a). Importantly, activation to red/yellow was
151 significantly stronger than blue in AMC, but not CLC/ALC (Supplementary Fig. 4b). The presence of the
152 bias for red/yellow in the global AMC fMRI signal shows that it is not due to selective sampling.

153 We also compared sharpness of hue tuning across the three patches, and found evidence for gradual
154 sharpening from CLC to ALC to AMC (Supplementary Fig. 5). Overall, the results so far show that (1) each
155 color patch contains a large population of hue-tuned cells, and (2) a transformation in hue tuning occurs
156 between CLC/ALC, where different hues are uniformly represented, and AMC, where red and yellow are
157 over-represented compared to green/blue.

158 **Representation of object shape by neurons in color patches**

159 Thus far, we have examined tuning to hue, averaged across object identity. However, since the color
160 patches have a stereotyped location relative to face patches, which represent facial shape, a natural
161 question is: how is object shape represented across color patches? To quantify the representation of
162 object shape, we computed responses to 82 objects averaged across 8 hues. MDS was conducted on
163 population responses of color-selective cells in three patches (Fig. 4a). All three patches displayed clear
164 grouping according to object category. But the amount of information about object shape was very
165 different between the three patches: the accuracy for identifying images, quantified by a nearest
166 neighbor classifier (see Methods), was significantly higher in CLC than ALC across all ten categories (Fig.
167 4b, $p<0.05$, 20000 iterations of random sampling with replacement, cell numbers were equalized to 55

168 for all three patches). Comparing ALC with AMC, identification accuracy was significantly higher in AMC
169 for humans and monkeys ($p < 0.05$) but not other categories (Fig. 4b). Comparing CLC with AMC, accuracy
170 was significantly higher for all categories ($p < 0.05$) except from monkeys ($p = 0.496$). A 2-way ANOVA
171 analysis to test significant interaction between area (2 levels) and category (10 levels) revealed
172 significant interactions for ALC and AMC ($F(9) = 2.03$, $p = 0.040$), CLC and AMC ($F(9) = 3.36$, $p = 0.001$), but
173 not ALC and CLC ($F(9) = 0.98$, $p = 0.459$).

174 Could the shape information observed in the color patches be a vestige of low-level shape selectivity
175 present in presumptive inputs to the color patch system, e.g., orientation-tuned cells in area V4? To
176 address this, we compared object shape representations in each color patch with those in two models,
177 AlexNet¹⁷ and HMAX¹⁸, a model for visual processing in V1-V4. The results show that the shape
178 representations in CLC, ALC and AMC are high-level, consistent with those in other parts of IT cortex
179 (Supplementary Fig. 6a-d). Note that here we only investigated representation of shape independent of
180 hue, thus the similarity between color patches and other regions in IT (Supplementary Fig. 6a) is
181 restricted to shape and does not indicate anything about color representation.

182 Within the face patch system, the most salient difference between patches is how they represent facial
183 identity across different views, with an increasingly view-invariant representation as one moves anterior
184 ⁹. Does a similar transformation in view-invariant object identity occur in the color patches? We
185 presented facial images of different identities at eight hues and three views: left/ right profiles and
186 frontal (Supplementary Fig. 2i). In AMC, we found cells mirror symmetrically-tuned to views (Fig. 4c, d
187 shows one example cell). The population response showed a correlation between responses to left and
188 right profile views of the same identity (Fig. 4e, f; $t(68) = -4.02$, $p = 1 * 10^{-4}$ between CLC and AMC; $t(82) = -$
189 3.48 , $p = 8 * 10^{-4}$ between ALC and AMC; $t(64) = -1.22$, $p = 0.23$ between CLC and ALC, Student's t-test),
190 similar to anterior face patch AL⁹. This mirror symmetric view invariance was weaker in CLC and ALC (Fig.
191 4e, f). This result shows that the representation of facial shape in AMC is not simply inherited from CLC.

192 **Co-representation of hue and shape by neurons in color patches**

193 The analyses so far have examined color and shape representations in isolation, and revealed that both
194 color and shape information are present in all three color patches. To gain a full picture of information
195 flow across patches, we next examined co-representation of the two information channels across
196 patches.

197 We first conducted MDS analyses on responses to all human face images in the stimulus set (Fig. 5a). We
198 found that the neural representation of colored faces differed between the three patches: although all
199 three patches showed grouping of images according to hue, this grouping was clearer in ALC and AMC
200 than CLC. Outside the color patches, images were grouped according to identity rather than hue. This
201 difference is illustrated by similarity matrices (Fig. 5b; for full similarity matrices see Supplementary Fig.

202 7): the 11*11 squares along the diagonal reflecting hue-specific representation are strongest in ALC,
203 followed by AMC, and least clear in CLC; the para-diagonal stripes indicating hue-invariant identity
204 information are only clearly observed in CLC but not the other two patches. Outside the color patches,
205 strong para-diagonal stripes are visible. To quantify relative contributions of hue and identity for each
206 category, we averaged correlation coefficients between population responses to images with the same
207 hue but different identity or same identity with different hue (Fig. 5c). We found comparable amounts
208 of hue and identity information for all 10 categories in CLC ($p=0.053, 0.359, 0.203, 0.128, 0.006, 0.025,$
209 $0.421, 0.001, 0.287, 0.329$, 20000 iterations of bootstrapping), but a clear bias for hue information in
210 ALC ($p=0, 0, 0, 0, 0, 0, 0.027, 0, 0, 0.004$). AMC was similar to ALC, showing a strong bias for hue, with
211 one prominent exception: for the category of monkey images, hue and identity information were
212 comparable ($p=0.34$). The presence of hue-invariant identity information about monkeys in AMC is
213 consistent with the superior ability to identify monkeys compared to other objects using AMC
214 population responses (Fig. 5b). The enhanced representation of monkey identities compared to other
215 objects in AMC adds support to our hypothesis that it is biased to represent colored faces and bodies.
216 Outside the color patches, there was significantly more information about identity compared to hue for
217 all the natural image categories (Fig. 5c).

218 We further analyzed the co-representation of hue information and category information using
219 responses averaged across all identities within one category. This revealed a bias for category
220 information in CLC and IT regions outside the color patches, and hue information in ALC and AMC (Fig.
221 5e, f). Examination of the time course of color information and identity information in all three patches
222 revealed that color was always faster than identity, even in CLC where identity information was slightly
223 stronger at the peak (Fig. 5d, g). To quantify the difference in temporal dynamics, we defined latency as
224 the first time point hue/shape information differed significantly from baseline ($p<0.01$, 20000 iterations
225 of bootstrapping, note that the each time point t indicates a time window [$t-25$ ms, $t+25$ ms]). In all
226 cases, color was faster than shape (shape identity vs. color: i.d.=50 ms and color=25 ms for CLC, i.d.=75
227 ms and color=50 ms for ALC, i.d.=75 ms and color=50 ms for AMC; shape category vs. color: category=50
228 ms and color=25 ms for CLC, category=75 ms and color=25 ms for ALC, category=75 ms and color=50 ms
229 for AMC).

230 Previous studies on object representation in IT from other labs employed linear classifiers to quantify
231 the amount of “linearly” decodable information from population response of neurons^{14 19}. This provides
232 a useful way to measure whether a particular dimension of information coded by a certain brain area is
233 “untangled” from other dimensions. We applied the same method to our data to quantify the amount of
234 shape-invariant hue information and hue-invariant shape information. Consistent with our analyses with
235 similarity matrices, we found that shape-invariant hue information was higher in anterior color patches
236 than in CLC, while hue-invariant shape information showed the opposite trend (Fig. 6). Thus suggests

237 that hue information does indeed become untangled from shape information along the color patch
238 pathway.

239 **Co-representation of hue and object shape at the single-cell level**

240 So far, we have shown that all three patches contain information about both object shape and hue.
241 What is the relative contribution of these two variables within single cells? Furthermore, what
242 integration rule is used by single cells in color patches to combine hue and shape information? To
243 answer these two questions, we performed a 2-way ANOVA analysis on single cell responses (Fig. 7a),
244 with 82 levels of shape and 8 levels of hue. A scatter plot of explained variance due to hue versus that
245 due to shape revealed an inverse relationship between the two, as expected (Fig. 7b1). Consistent with
246 previous population analyses, AMC and ALC neurons are more biased to hue than CLC neurons (Fig. 7b2):
247 ALC and AMC are significantly more hue-biased than CLC ($t(118)=-7.61$, $p=7*10^{-12}$ and $t(143)=-9.0$,
248 $p=2*10^{-15}$ respectively, Student's t-test); CLC is significantly more hue-biased than outside ($t(96)=-6.1$,
249 $p=3*10^{-8}$); ALC and AMC are not significantly different ($t(133)=0.2$, $p=0.85$). However, it is worth noting
250 that all three color patches did carry a significant amount of shape information. In CLC, the mean
251 amount of variance accounted for by shape was 37.4%. In ALC, the mean amount of variance accounted
252 for by shape was 18.8%, while in AMC, it was 20.2% (Fig. 7b2; note that given noise in the data, the
253 relative contribution from shape is over-estimated, since there are more parameters for shape (82) than
254 for hue (8)). We performed similar analysis, but using 10 coarse shape categories or shape-within-
255 categories (on average 8.2 shapes in each category) to define the shape variables. We found that in
256 these two cases, AMC and ALC neurons, but not CLC neurons, were clearly hue biased (Fig. 7b3-b4).
257 Comparing four regions for coarse shape categories: ALC and AMC are significantly more hue-biased
258 than CLC ($t(118)=-4.5$, $p=1*10^{-5}$ and $t(143)=-4.6$, $p=8*10^{-6}$ respectively, Student's t-test); CLC is
259 significantly more hue-biased than outside ($t(96)=-7.9$, $p=4*10^{-12}$); ALC and AMC are not significantly
260 different ($t(133)=0.6$, $p=0.5$). Comparing four regions for shape-within-categories: ALC and AMC are
261 significantly more hue-biased than CLC ($t(118)=9.0$, $p=3*10^{-15}$ and $t(143)=9.5$, $p=3*10^{-17}$ respectively,
262 Student's t-test); CLC is significantly more hue-biased than outside ($t(96)=-6.5$, $p=3*10^{-9}$); ALC and AMC
263 are not significantly different ($t(133)=-0.8$, $p=0.4$). As expected, all color patch neurons show a
264 significant main effect for hue (ANOVA, $p<0.001$), while most color patch neurons (with the exception of
265 2 ALC neurons and 1 AMC neuron) showed a significant main effect for shape (Fig. 7b5). Finally, we
266 found a large portion of color patch neurons showed a significant interaction between hue and shape
267 (41/65 (=63%) CLC cells, 40/55 (=73%) ALC cells, 47/80 (=59%) AMC cells, and 0/33 cells outside color
268 patches, Fig. 7b6).

269 The nonlinear interaction between shape and hue is further supported by comparison of MDS analysis
270 of shape responses at the best and worst hues. If cells were linearly adding the two variables, then the
271 MDS plots at the two hues should be identical; however, we found that shape representation at the

272 worst hue was compressed compared to that at the best hue (Fig. 7c1, c2; $t(96)=6.4$ and $p=7*10^{-9}$
273 between outside vs. CLC; $t(86)=11.8$ and $p=1*10^{-15}$ between outside and ALC; $t(111)=9.0$ and $p=6*10^{-15}$
274 between outside and AMC, Student's t-test). Overall, these results suggest that cells in color patches are
275 not simply summing shape and color inputs, but are nonlinearly combining the two.

276 The presence of nonlinear interaction between shape and hue raises the question whether the main
277 effects for hue and shape are real, i.e. the main effects may appear only in some conditions but not at all
278 in others. To address this, for each cell in each patch, we quantified hue tuning for each shape by
279 computing the variance of responses across hue to the shape, and selected 27 shapes with the best and
280 worst hue tuning. Similarly, we selected 3 hues with the best and worst shape tuning for each cell. We
281 computed the correlation between hue tuning for the "best" and the "worst" shapes, and did the same
282 for shape tuning. We found all cells except 1 ALC cell showed positive correlation between hue tuning
283 for the "best" and the "worst" shapes, and most cells (65/65 CLC cells, 51/55 ALC cells, 71/80 AMC cells)
284 showed positive correlation between shape tuning for the "best" and the "worst" hues. Furthermore,
285 we performed ANOVA on the "worst" shapes/hues, and a high percentage of cells showed significant
286 ($p<0.001$) main effects for hue (34/65 CLC cells, 42/55 ALC cells, 69/80 AMC cells)/shape (64/65 CLC cells,
287 47/55 ALC cells, 69/80 AMC cells). We note that different selection criterion was applied to color patches
288 and neurons outside the color patches in the above two analyses (Figs. 5 and 7): while only color
289 selective cells were used for color patches, all cells were pooled for "outside". However, all the results
290 comparing color patches to outside the color patches remained consistent when we pooled all cells in
291 color patches.

292

293 Discussion

294 In this study, we found that three macaque color patches, CLC, ALC, and AMC, all encode significant
295 information about both hue and object identity. Two clear transformations occur across the three
296 patches. The first transformation, from CLC to ALC, reduces information about object identity. The
297 second transformation, from ALC to AMC, mainly affects representation of hue: color space is
298 represented in a dramatically distorted way in AMC, with over-representation of yellow and red, the
299 natural colors of mammal faces and bodies. Furthermore, AMC develops an expanded representation of
300 primate faces compared to other categories, displaying hue-invariant representation of monkey identity.

301 Our study broadens our conception of the function of IT cortex. A generally accepted notion is that the
302 purpose of IT is to represent object identity invariant to accidental changes, and this is achieved through
303 a hierarchy culminating in cells in anterior IT tuned to object identity invariant to accidental changes
304 ^{9,18,20}. For example, in the face patches, tuning to facial identity becomes more invariant to view going
305 from ML/MF to AL to AM ⁹. Applying this principle to colored objects, one might expect to find a

306 sequence of areas tuned to object hue and shape combinations that show increasing invariance to
307 accidental transformations in view, lighting, etc. (Fig. 8b). Instead, we found the existence of a
308 specialized network in which shape information *decreases* along the IT hierarchy, while hue information
309 *increases* (Fig. 8a). Since shape is one of the strongest cues to object identity, this calls into question the
310 current picture of IT as a monolithic hierarchical feedforward network for computing object identity²⁰.
311 IT appears to generate an array of high-level representations of objects that can facilitate different
312 object-related tasks, including the fundamental task of identifying the color of an object.

313 What is the role of AMC? Why should there exist an area with neurons that represent hue irrespective
314 of shape, but mainly for red and yellowish things, but then also the shape of faces irrespective of hue?
315 At first glance, this seems unparsimonious. One possible explanation is that AMC provides an important
316 intermediate link between a multi-purpose shape-invariant hue representation and a representation
317 specialized for the color of animal faces/bodies. The diversity in responses to different faces in AMC
318 could compensate for the relative homogeneity in their responses to hues, exploiting the coding space
319 previously occupied (in ALC) by bluish hues. This would facilitate the wiring of classifiers trained to
320 identify faces using both shape and color cues. Future experiments exploring responses in color patches
321 during performance of active tasks including face and color categorization may shed further light on the
322 functional role of each patch.

323 Few previous studies have examined color tuning in IT cortex. Most have reported aggregate statistics
324 based on random sampling of IT cortex²¹⁻²⁵, and have come to conflicting conclusions regarding the
325 prevalence of hue-tuned cells, from 15%²¹ to 69%²², as well as the extent to which shape and hue
326 interact in single cells; Komatsu and Ideura²² reported no interaction, while Edwards et al.²³ reported
327 strong nonlinear interaction. One pioneering study of color transformations in IT reported two clusters
328 of color-selective cells in IT cortex, one in posterior and one in anterior IT, and showed a difference
329 between these two regions in their color tuning as a function of luminance¹². Our study shows that
330 there are at least three clusters of color-selective cells in IT that are strongly anatomically connected.
331 Most importantly, our study demonstrates the importance of (1) studying the *co-representation* of color
332 and object shape within each color patch, and (2) studying multiple patches within the IT color network
333 using the same stimuli. Only by taking both of these steps, could we reveal the transformations in the
334 brain's representation of colored objects for the first time.

335

336

337

338

339

340

341

342

343

344

345

346

347

348 **Methods**

349

350 **Color Patch Localization**

351

352 All procedures conformed to local and US National Institutes of Health guidelines, including the US
353 National Institutes of Health Guide for Care and Use of Laboratory Animals. All experiments were
354 performed with the approval of the Caltech Institutional Animal Care and Use Committee (IACUC).

355

356 Four male rhesus macaques were trained to maintain fixation on a small spot for juice reward (three of
357 the animals were used for color patch recordings, while the fourth animal was used solely for control
358 recordings outside color patches). Monkeys were scanned in a 3T TIM (Siemens, Munich, Germany)
359 magnet equipped with AC88 gradient insert while passively viewing images on a screen. Feraheme
360 contrast agent was injected to improve signal/noise ratio. Color patches were determined by identifying
361 regions responding significantly more to moving equiluminant red/green color gratings (2.9 cycles per
362 degree, drifting 0.75 cycles per s) than moving black-white gratings, the same stimuli as a previous study
363 ⁵, and were confirmed across multiple independent scan sessions. In monkey M1, color patches CLC, ALC
364 and AMC were found bilaterally. In monkey M2, color patches CLC and ALC were found bilaterally; color
365 patches AMC and AFC were found unilaterally in the left hemisphere (AMC was found by
366 microstimulating ALC in the same hemisphere, see below). In monkey M3, color patches CLC and ALC
367 were found bilaterally; color patch AFC was found unilaterally in the left hemisphere.

368

369 **Microstimulation**

370

371 To reveal the anatomical connectivity of color patches and to localize the most anterior color patch AMC
372 in monkey M2, we stimulated ALC¹⁵. The stimulation protocol followed a block design. We normally
373 interleaved 9 blocks of fixation-only with 8 blocks of fixation plus electrical microstimulation; we always
374 started and ended with a fixation-only block. During microstimulation blocks we applied one pulse train
375 per second, lasting 200 ms with a pulse frequency of 300 Hz. Bipolar current pulses were charge
376 balanced, with a phase duration of 300 μ s and a distance between the two phases of 150 μ s. We used a
377 current amplitude of 300 μ A. Stimulation pulses were delivered with a computer-triggered pulse
378 generator (S88X; Grass Technologies) connected to a stimulus isolator (A365; World Precision
379 Instruments), which interfaced with different and indifferent electrodes through a coaxial cable. All
380 stimulus generation equipment was stored in the scanner control room; the coaxial cable was passed
381 through a wave guide into the scanner room. We performed electrophysiological recording at the site of
382 stimulation immediately prior to stimulation, to confirm correct electrode placement, as revealed by a
383 high number of hue-selective units.

384

385 **Single-unit recording**

386

387 Tungsten electrodes (18–20 Mohm at 1 kHz, FHC) were back loaded into plastic guide tubes. Guide
388 tubes length was set to reach approximately 3–5 mm below the dura surface. The electrode was
389 advanced slowly with a manual advancer (Narishige Scientific Instrument, Tokyo, Japan). Neural signals
390 were amplified and extracellular action potentials were isolated using the box method in an on-line
391 spike sorting system (Plexon, Dallas, TX, USA). Spikes were sampled at 40 kHz. All spike data was re-
392 sorted with off-line spike sorting clustering algorithms (Plexon). Only well-isolated units were considered
393 for further analysis. We targeted patches CLC (n=35; right hemisphere) and AMC (n=24; right
394 hemisphere) in monkey M1, patches CLC (n=52; right hemisphere), ALC (n=43; left hemisphere) and
395 AMC (n=66; left hemisphere) in monkey M2, and patch ALC (n=22; right hemisphere) in monkey M3 for
396 single-unit recordings. In addition, we targeted face patch AM (n = 10; right hemisphere) and a region of
397 anterior IT on the ventral bank of the inferotemporal gyrus outside the color patches in M1 (n=10; right
398 hemisphere), and a region of middle IT on the ventral bank of superior temporal sulcus outside the color
399 patches in M4 (n = 13; left hemisphere). Electrodes were lowered through custom angled grids that
400 allowed us to reach the desired targets; custom software was used to design the grids and plan the
401 electrode trajectories²⁶. For each patch, results were qualitatively the same across different monkeys
402 and therefore were pooled together for population analyses. Multiple different tracks were used to
403 target each patch; in particular, for AMC, we designed ten distinct approach angles using different grids
404 to ensure even sampling.

405

406 **Behavioral Task and Visual Stimuli**

407
408 Monkeys were head fixed and passively viewed the screen in a dark room. Stimuli were presented on a
409 CRT monitor (DELL P1130). The intensity of the screen was measured using a colorimeter (PR650, Photo
410 Research) and linearized for visual stimulation. Screen size covered 27.7*36.9 visual degrees and
411 stimulus size spanned 5.7 degrees. The fixation spot size was 0.2 degrees in diameter. Images were
412 presented in random order using custom software. Eye position was monitored using an infrared eye
413 tracking system (ISCAN). Juice reward was delivered every 2–4 s if fixation was properly maintained.

414
415 For visual stimulation, all images were presented for 200 ms interleaved by 200 ms of a gray screen.
416 Each image was presented 7–10 times to obtain reliable firing rate statistics. In this study, 2 different
417 stimulus sets were used:

- 418 1) A set of 82 images of 10 different categories, varied in 8 different hues. Original images and
419 grayscale images with the same luminance profile were also presented (Supplementary Fig.
420 2a, for details see below).
- 421 2) A set of 33 human face images (Supplementary Fig. 2i), 3 different views of 11 identities,
422 varied in 8 different hues.

423 424 **Color stimulus generation**

425
426 For our color stimuli, we started with a set of 55 object images collected from the internet, and 11
427 frontal human faces from an on-line database (FEI face database:
428 <http://fei.edu.br/~cet/facedatabase.html>). We transformed the color of each image in the following way:
429 For a given pixel with RGB value (r,g,b) , its chromaticity coordinates and luminance (u,v,L) , CIE 1960
430 were estimated by first computing the frequency spectrum of the pixel by summing the frequency
431 spectra for r , g , and b , each measured separately using a spectrophotometer (PR650, Photo Research),
432 and then converted into chromaticity coordinates (<http://www.cvrl.org>). The mean luminance of each
433 image was equalized to the background (2.9 cd/m^2). We then computed the distance of chromatic
434 coordinates of each pixel to “white” ($u=0.2105, v=0.3158$, filled circle in Fig. 1e). Eight different colors
435 with the same distance but varying angles (open circles in Fig. 1e, starting from 0° , going clockwise at 45°
436 step) were then computed and converted back into an RGB value keeping the luminance (L) unchanged.
437 Repeating this for every pixel resulted in 8 images of pure hues (Fig. 1e right). A grayscale image with
438 same luminance, but “white” chromatic coordinate, was also generated. The stimulus set included
439 simple geometric patterns (8th and 9th row in Supplementary Fig. 2a). For these images, we set the hue
440 of the original image to orange, with mean luminance equal to background. We also included a category
441 of phase scrambled images (last row in Supplementary Fig. 2a). Eight images from the first nine
442 categories were randomly selected and phase-scrambled, keeping the relative phase between different
443 cone components constant.

444 The other stimulus set was generated in the same way, but using only face images with different head
445 orientations (left profile, frontal and right profile, Supplementary Fig. 2i).

446 **Color selectivity**

447 For each repetition of the stimulus set, responses to each of the eight hues were averaged across 82
448 objects. Classical analysis of variance (ANOVA) was performed to test the statistical significance of the
449 differences among eight hue groups, each group containing multiple repetitions of the stimulus set. Only
450 significantly tuned cells ($p < 0.001$) were used for further analysis.

451 **Multi-dimensional scaling**

452 The number of spikes in a time window of 50-350 ms after stimulus onset was counted for each stimulus.
453 The responses of each cell to all stimulus conditions were normalized to 0 mean and unit variance. To
454 study the neural representation of a single feature, such as hue (Fig. 3b) or object shape (Fig. 4a),
455 responses were averaged across the irrelevant feature (shape in Fig. 3b and hue in Fig. 4a). Classical
456 multi-dimensional scaling was performed on the population responses in each patch, using a Euclidean
457 distance metric and the MATLAB command `cmdscale`.

458 **Neural distance**

459 The responses of each cell to all stimulus conditions were normalized to 0 mean and unit variance.
460 Euclidean distance between the normalized population responses to two stimulus conditions was used
461 to quantify the “neural” distance between these two conditions (Fig. 3c).

462 **Similarity matrix**

463 Based on the same normalized population response, an $n \times n$ similarity matrix of correlation coefficients
464 was computed between the population response vectors (across all color-selective cells, averaged over
465 stimulus repeats) to each of the n interested conditions.

466 **Decoding analysis**

467 To quantify the amount of information about object shape in all three patches, we trained a nearest
468 neighbor classifier: the population response for one particular object averaged across hues in two thirds
469 of the trials was used to define a “template” response for that object. For testing, the population
470 response to one image averaged across the remaining one third of the trials (but not across hues) was
471 compared to each of the 82 “templates”, and the object “template” with minimal distance to actual
472 response was defined as the output of the classifier.

473 We also employed SVM decoding models as in previous papers^{14,19}. In brief, we randomly selected a
474 number of units from each area, and trained an SVM model for each selection to decode hue

475 information independent of shape or shape information independent of hue using “one vs. rest”
476 approach. We used half the trials to train the SVM model and the remaining half to validate the model.
477 The results shown are validated accuracies.

478 **Convolutional neural network modeling**

479
480 To investigate shape representation in three color patches, we loaded 82 objects with 8 different hues
481 into two pre-trained neural networks: 1) a matlab implementation of Alexnet ¹⁷ :
482 <http://www.vlfeat.org/matconvnet/pretrained/>. This network contains 21 layers: 1st, 5th, 9th, 11th and
483 13th layers are the outputs of convolution units; 2nd, 6th, 10th, 12th and 14th, 17th and 19th layers are the
484 results of rectification; 3rd and 7th layers are the results of normalization; 4th, 8th, 15th layers are the
485 results of max pooling; 16th, 18th and 20th layers are fully connected layers; 21st layer is the output layer
486 (softmax). This network has been pre-trained to identify a thousand objects.
487 2) a matlab implementation of HMAX model ¹⁸: <http://maxlab.neuro.georgetown.edu/hmax.html#code>.
488 This network implements the basic architecture of the HMAX model (S1, C1, S2, C2) and has been pre-
489 trained with a set of random natural images. Activations of each unit to each object were averaged
490 across hues to analyze the representation for shape alone. For the case of the HMAX model, since the
491 network only allows grayscale images as input, we presented the grayscale version of the 82 objects to
492 the network to analyze shape representation.

493 **Population statistics**

494 To determine statistical significance for parameters estimated using population responses, such as
495 correlation, a bootstrap method was employed: neurons were randomly sampled from the population
496 with replacement; 20000 bootstrap samples with equal numbers of neurons were created. A population
497 statistic was computed for each bootstrap sample. The P-value of the null hypothesis was determined by
498 comparing population statistics from 20000 iterations of bootstrapping.

499

500

501

502

503

504

505

506

507 **References:**

- 508 1 Livingstone, M. S. & Hubel, D. H. Anatomy and physiology of a color system in the primate visual
509 cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **4**, 309-
510 356 (1984).
- 511 2 Conway, B. R. Spatial structure of cone inputs to color cells in alert macaque primary visual
512 cortex (V-1). *The Journal of neuroscience : the official journal of the Society for Neuroscience* **21**,
513 2768-2783 (2001).
- 514 3 Lim, H., Wang, Y., Xiao, Y., Hu, M. & Felleman, D. J. Organization of hue selectivity in macaque
515 V2 thin stripes. *Journal of neurophysiology* **102**, 2603-2615, doi:10.1152/jn.91255.2008 (2009).
- 516 4 Conway, B. R., Moeller, S. & Tsao, D. Y. Specialized color modules in macaque extrastriate cortex.
517 *Neuron* **56**, 560-573, doi:10.1016/j.neuron.2007.10.008 (2007).
- 518 5 Lafer-Sousa, R. & Conway, B. R. Parallel, multi-stage processing of colors, faces and shapes in
519 macaque inferior temporal cortex. *Nature neuroscience* **16**, 1870-1878, doi:10.1038/nn.3555
520 (2013).
- 521 6 Tanaka, K. Inferotemporal cortex and object vision. *Annual review of neuroscience* **19**, 109-139,
522 doi:10.1146/annurev.ne.19.030196.000545 (1996).
- 523 7 Brincat, S. L. & Connor, C. E. Underlying principles of visual shape selectivity in posterior
524 inferotemporal cortex. *Nature neuroscience* **7**, 880-886, doi:10.1038/nn1278 (2004).
- 525 8 Tsao, D. Y., Moeller, S. & Freiwald, W. A. Comparing face patch systems in macaques and
526 humans. *Proceedings of the National Academy of Sciences of the United States of America* **105**,
527 19514-19519, doi:10.1073/pnas.0809662105 (2008).
- 528 9 Freiwald, W. A. & Tsao, D. Y. Functional compartmentalization and viewpoint generalization
529 within the macaque face-processing system. *Science* **330**, 845-851,
530 doi:10.1126/science.1194908 (2010).
- 531 10 Field, G. D. *et al.* Functional connectivity in the retina at the resolution of photoreceptors.
532 *Nature* **467**, 673-677, doi:10.1038/nature09424 (2010).
- 533 11 Horwitz, G. D. & Hass, C. A. Nonlinear analysis of macaque V1 color tuning reveals cardinal
534 directions for cortical color processing. *Nature neuroscience* **15**, 913-919, doi:10.1038/nn.3105
535 (2012).
- 536 12 Namima, T., Yasuda, M., Banno, T., Okazawa, G. & Komatsu, H. Effects of luminance contrast on
537 the color selectivity of neurons in the macaque area v4 and inferior temporal cortex. *The Journal*
538 *of neuroscience : the official journal of the Society for Neuroscience* **34**, 14934-14947,
539 doi:10.1523/JNEUROSCI.2289-14.2014 (2014).
- 540 13 Tsao, D. Y., Freiwald, W. A., Tootell, R. B. & Livingstone, M. S. A cortical region consisting entirely
541 of face-selective cells. *Science* **311**, 670-674, doi:10.1126/science.1119983 (2006).

- 542 14 Hung, C. P., Kreiman, G., Poggio, T. & DiCarlo, J. J. Fast readout of object identity from macaque
543 inferior temporal cortex. *Science* **310**, 863-866, doi:10.1126/science.1117593 (2005).
- 544 15 Moeller, S., Freiwald, W. A. & Tsao, D. Y. Patches with links: a unified system for processing faces
545 in the macaque temporal lobe. *Science* **320**, 1355-1359, doi:10.1126/science.1157436 (2008).
- 546 16 Kornblith, S., Cheng, X., Ohayon, S. & Tsao, D. Y. A network for scene processing in the macaque
547 temporal lobe. *Neuron* **79**, 766-781, doi:10.1016/j.neuron.2013.06.015 (2013).
- 548 17 Krizhevsky, A., Sutskever, I. & Hinton, G. ImageNet classification with deep convolutional neural
549 networks. *In Proc. Advances in Neural Information Processing Systems 25*, 1090-1098 (2012).
- 550 18 Riesenhuber, M. & Poggio, T. Hierarchical models of object recognition in cortex. *Nature*
551 *neuroscience* **2**, 1019-1025, doi:10.1038/14819 (1999).
- 552 19 Pagan, M., Urban, L. S., Wohl, M. P. & Rust, N. C. Signals in inferotemporal and perirhinal cortex
553 suggest an untangling of visual target information. *Nature neuroscience* **16**, 1132-1139,
554 doi:10.1038/nn.3433 (2013).
- 555 20 Yamins, D. L. *et al.* Performance-optimized hierarchical models predict neural responses in
556 higher visual cortex. *Proceedings of the National Academy of Sciences of the United States of*
557 *America* **111**, 8619-8624, doi:10.1073/pnas.1403112111 (2014).
- 558 21 Desimone, R., Albright, T. D., Gross, C. G. & Bruce, C. Stimulus-selective properties of inferior
559 temporal neurons in the macaque. *The Journal of neuroscience : the official journal of the*
560 *Society for Neuroscience* **4**, 2051-2062 (1984).
- 561 22 Komatsu, H. & Ideura, Y. Relationships between color, shape, and pattern selectivities of
562 neurons in the inferior temporal cortex of the monkey. *Journal of neurophysiology* **70**, 677-694
563 (1993).
- 564 23 Edwards, R., Xiao, D., Keyser, C., Foldiak, P. & Perrett, D. Color sensitivity of cells responsive to
565 complex stimuli in the temporal cortex. *Journal of neurophysiology* **90**, 1245-1256,
566 doi:10.1152/jn.00524.2002 (2003).
- 567 24 McMahan, D. B. & Olson, C. R. Linearly additive shape and color signals in monkey
568 inferotemporal cortex. *Journal of neurophysiology* **101**, 1867-1875, doi:10.1152/jn.90650.2008
569 (2009).
- 570 25 Yasuda, M., Banno, T. & Komatsu, H. Color selectivity of neurons in the posterior inferior
571 temporal cortex of the macaque monkey. *Cerebral cortex* **20**, 1630-1646,
572 doi:10.1093/cercor/bhp227 (2010).
- 573 26 Ohayon, S. & Tsao, D. Y. MR-guided stereotactic navigation. *Journal of neuroscience methods*
574 **204**, 389-397, doi:10.1016/j.jneumeth.2011.11.031 (2012).

575

576

577 **Acknowledgments**

578

579 This work was supported by the Howard Hughes Medical Institute, the Tianqiao and Chrissy Chen
580 Institute for Neuroscience at Caltech, and the Swartz Foundation (fellowship to LC). We thank Nicole
581 Schweers for technical support, members of the Tsao lab, Margaret Livingstone, and Simon Kornblith for
582 critical comments.

583

584 **Author contributions**

585

586 L.C. and D.Y.T. designed the experiments, interpreted the data, and wrote the paper. L.C. and P.B.
587 conducted the experiments and analyzed the data.

588

589 **Additional information**

590 **Competing interests:** The authors declare no competing financial interests.

591

592 **Figure legends:**

593 **Figure 1. Recording sites, connectivity, and color stimuli.**

594 **a**, To identify the correct color of an object, e.g. an apple (*left*), local hue information (*right*) needs to
595 be integrated with global shape information. **b**, Schemes for co-representing color and object shape
596 information in visual system. Initially, before the visual system has explicitly segmented objects, color
597 information and object shape information are largely entangled, with individual cells participating in the
598 coding of multiple hues and object shapes. Two main strategies could be used to represent colored-
599 objects in an organized way: 1) segregation of color and object information into parallel channels,
600 resulting in object shape-invariant color-selective units and color-invariant shape selective units (top); 2)
601 formation of units sharply tuned to both color and object shape (bottom). **c**, Coronal and Sagittal slices
602 showing location of fMRI-identified face (blue) and color patches (yellow) in one monkey (M2) targeted
603 for recording; dark black line indicates electrode. The most anterior color patch was not observed with
604 fMRI using the color localizer in this animal, and was located by electrical microstimulation in ALC
605 (bottom panel, changes in BOLD signal of the identified voxels during microstimulation shown below).
606 Source: Tsao lab. **d**, Comparison between color patches identified by color localizer (top) and by
607 microstimulation (bottom). The contrasts are overlaid on high-resolution coronal slices. Asterisk (*)
608 indicates the stimulation site (ALC). The anterior-posterior position of each slice in mm relative to the
609 interaural line is given in the top right corner. **e**, 82 images of 10 categories were used (see

610 Supplementary Fig. 2a for all the stimuli). Each image underwent a series of transformation in hues. For
611 each pixel of the image, luminance was kept constant, while chromatic coordinates (CIE 1960) fell on a
612 circle with the same distance to “white” (filled circle) as the original pixel. Eight hues with different
613 angles were used (open circles, starting from 0°, going clockwise at 45° step). A grayscale image with the
614 same luminance and the natural color image were also presented.

615 **Figure 2. Responses of color patch neurons to the color stimuli**

616 Responses of all neurons in three patches to all grayscale images, colored human faces, monkey faces
617 and magic cubes, sorted according to hue preference of the average response across all 82 stimuli.
618 Color-selective cells and non-selective cells are shown separately. Responses of IT cells outside color
619 patches and cells in face patch AM are also shown. For each cell, baseline was subtracted and the
620 response was normalized.

621 **Figure 3. Representation of color by color-selective neurons in color patches.**

622 **a**, Responses of all color-selective neurons averaged across stimuli within each category to images of 8
623 different hues, together with gray (left-most column) and natural color (right most-column), sorted in
624 this same way as Fig. 2. **b**, Neural representation of colors in the activities of color-selective neurons.
625 Shown are two-dimensional plots of the results of multi-dimensional scaling (MDS) analyses conducted
626 for neurons in three color patches. Responses to each color condition were averaged across all mammal
627 images (humans faces, monkey faces, and mammal bodies; these were selected because color tuning
628 was most consistent between these three categories across all three patches, see Supplementary Fig.
629 3b). Original color is indicated by a disk of mixed color. **c**, Neural distances of each hue to its two
630 neighboring hues, for all three patches, computed using population responses of color-selective cells.
631 Error bars represent s.d. of 20000 iterations of bootstrapping. Inhomogeneity was quantified by
632 computing the ratio between the s.d. of the 8 bars and the mean of the 8 bars: 0.21±0.02 for CLC;
633 0.12±0.02 for ALC; 0.35±0.02 for AMC ($p < 0.001$ between CLC and AMC; $p < 0.001$ between ALC and AMC;
634 $p = 0.0103$ between CLC and ALC, 20000 iterations of bootstrapping, see Methods). **d**, Population
635 similarity matrices of 10 color conditions in three color patches. A 10*10 matrix of correlation
636 coefficients was computed between responses of all color-selective neurons averaged across objects. **e**,
637 For five different types of objects: grapes, watermelon, birds, gratings and rubik’s cube, the number of
638 AMC cells preferring each of the eight hues was counted. In all five cases, the distribution was
639 significantly different from homogeneity (chi-square test: $p < 0.001$; $\chi^2(7) = 26.3, 28.6, 31.0, 38.6$ and 28.4
640 respectively).

641 **Figure 4. Representation of object shape by color-selective neurons in color patches.**

642 **a**, Neural representation of object shapes in the activities of color-selective neurons for three patches,
643 shown as two-dimensional MDS plots. Responses to each object shape were averaged across 8 hues. **b**,
644 Decoding accuracies for identifying one object out of 82 objects based on population responses in three
645 color patches, averaged across object identities within each category (see Methods). Error bars
646 represent s.e. Dashed lines indicate chance level ($1/82=1.2\%$). **c**, Raster plot showing responses of an
647 AMC neuron to colored human faces of 11 identities at three views: frontal, left- and right-profile
648 (Supplementary Fig. 2i). **d**, The same response in (c) averaged across 8 hues, showing strong correlation
649 between left- and right-profiles, but not between frontal and profile views. **e**, Correlation between
650 responses to left and right profile views was computed across identities for each cell. Mean and s.e. of
651 all neurons in three patches are plotted ($n=26$ CLC cells; $n=40$ ALC cells; $n=44$ AMC cells). Student's t-test
652 was used to determine statistical significance between patches ($*=p<0.05$, $**=p<0.01$). **f**, Population
653 similarity matrices of 11 identities*3 views in three color patches. The paradiagonal stripes in AMC
654 indicate high correlation between responses to mirror-symmetric views of the same identity (red
655 arrows).

656 **Figure 5. Co-representation of hue and object identity in color patches.**

657 **a**, Comparison of MDS plots of responses to all human face images in all three color patches and outside
658 color patches. For clarity, the original natural color images are not shown. In ALC and AMC, images were
659 clearly grouped according to hue, while in CLC this grouping is less clear. Outside the color patches,
660 images were grouped according to identity, but not hue. **b**, Population similarity matrices computed
661 from responses to human face images in three color patches and outside color patches. Correlation
662 coefficients were computed between responses to 11 identities and 8 hues. **c**, Hue information and
663 identity information for images of 10 categories in three color patches and outside color patches. Hue
664 information was quantified as the mean correlation between responses to images with the same hue
665 but different identity within the same category, while identity information was quantified as the mean
666 correlation between responses to images of the same identity with different hues. Note that here we
667 are quantifying shape-invariant hue tuning, which will be affected by both shape tuning and color tuning;
668 in particular, cells with strong shape tuning will show low shape-invariant hue tuning, even if they have
669 perfectly consistent hue tuning across shapes. Error bars represent s.d. of 2000 iterations of
670 bootstrapping. Statistical significance was determined between hue and identity information for each
671 category in three color patches and outside color patches ($*=p<0.05$, $**=p<0.01$). **d**, Amplitude of hue
672 and identity information for three patches and outside color patches, computed over a 50 ms sliding
673 time window, were averaged across all 10 categories. Shaded regions indicate s.d. estimated by 2000
674 iterations of bootstrapping. **e**, Co-representation of hue and category in all three patches and outside
675 color patches. Responses of each cell were averaged across different identities within a category. A
676 matrix of correlation coefficients was computed between responses to 10 categories* 8 hues. **f and g**,
677 Same as (c) and (d), but for hue and category information quantified by matrices in (e), $**=p<0.01$.

678 **Figure 6. Decoding shape-invariant color and color-invariant shape from color patches.**

679
680 **a1**, Svm models were trained to classify hues independent of shape. Population response of different
681 number of randomly selected units was used as the input to the model. Half of the trials were used for
682 training and the rest half for cross-validation. Shape-invariant hue could be significantly better decoded
683 by AMC and ALC population than than CLC (for 50 units, $p < 0.01$ for both comparisons, 2000 iterations of
684 bootstrapping). Furthermore, ALC shows better overall decoding than AMC ($p = 0.013$). Dashed line
685 indicates chance level ($1/8 = 12.5\%$). Results are averages across 2000 iterations of random sampling.
686 Errorbars represent s.d. **a2**, similar to (a1), but only quantifies decoding accuracy for two hue categories:
687 red and yellow. Decoding based on AMC is better than ALC, but not significant ($p = 0.185$). **a3**, similar to
688 (a1), but for a combined population of anterior color patch neurons. **b**, similar to (a1), but for hue-
689 invariant shape decoding. CLC is significantly better than ALC and AMC (for 50 units, $p < 0.01$ between
690 CLC and ALC, $p = 0.028$ between CLC and AMC). Furthermore, neurons outside color patches shows
691 better performance than color patch neurons, but only significantly better than ALC and AMC (For 25
692 units, $p < 0.01$ between outside and ALC, $p = 0.019$ between outside and AMC and $p = 0.186$ between
693 outside and CLC). Dashed line indicates chance level ($1/82 = 1.2\%$).

694

695 **Figure 7. Analysis of single-cell responses.**

696 **a**, Responses of 12 example neurons to the full stimulus set. Each row represents one color condition,
697 and each column represents one object shape. **b**, 2-way ANOVA analysis examining main effects of
698 shape and hue, as well as interactions. 2-way ANOVA analysis with 8 levels of hue and 82 levels of shape
699 was performed on responses of each individual neuron. **b1**, Relationship between explained variances
700 by two main effects for all neurons. Lines represent linear fits to cells in each patch. **b2**, Distribution of
701 shape preference in all three patches and outside the color patches, defined by the explained variance
702 by shape divided by the sum of both main effects. Arrows indicate population averages. **b3**, similar to
703 (b2), but using only coarse shape categories as shape variables. **b4**, similar to (b2), but using only fine
704 shapes within each shape category as shape variables. ANOVA analysis was carried out for each shape
705 category independently, with 8 levels of hue and n levels of shape ($n = \text{number of shapes within this}$
706 shape category). For each neuron, shape preference was computed and averaged across categories. **b5**,
707 For 2-way ANOVA with 8 levels of hue and 82 levels of shape, F-values for both main effects are plotted
708 against each other in log-scale. Gray lines indicate significance level ($p = 0.001$). **b6**, Distribution of F-
709 values for the interaction between hue and shape. Gray dashed lines indicate significance level
710 ($p = 0.001$). **c1**, For each cell in ALC, we determined the best hue and the worst hue based on responses
711 to 8 hues averaged across 82 object shapes. MDS analyses were conducted on shape responses for the
712 best hue or the worst hue of each cell. Two MDS plots are shown at the same scale. **c2**, For each cell, the
713 ratio between standard deviations of shape responses at the worst hue and the best hue was computed.

714 If the cells were linearly adding hue and shape, the two standard deviations should be identical.
715 Therefore the ratio between these two reflects the extent of nonlinearity in the interaction of hue and
716 shape. **= $p < 0.01$, Student's t-test.

717

718 **Figure 8. Extending the conventional view of IT: a theory of color processing in IT.**

719 **a**, Schematic summary of the co-representation of hue and object shape in three color patches. Here,
720 each oval represents the receptive field of one “idealized” color neuron in the 2-d object space spanned
721 by hue and object shape. **b**, Conventional view of IT predicts that the major transformation of colored-
722 object representation from posterior to anterior IT is the generation of invariance to accidental changes
723 (e.g., view). Here each ellipsoid represents the receptive field of one “idealized” neuron in the 3-d object
724 space spanned by hue, shape, and view. For both posterior and anterior IT, two dimensional slices at a
725 fixed “view” should look the same as the schematic for CLC (a1).

bioRxiv preprint first posted online Oct. 19, 2017; doi: <http://dx.doi.org/10.1101/205104>. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

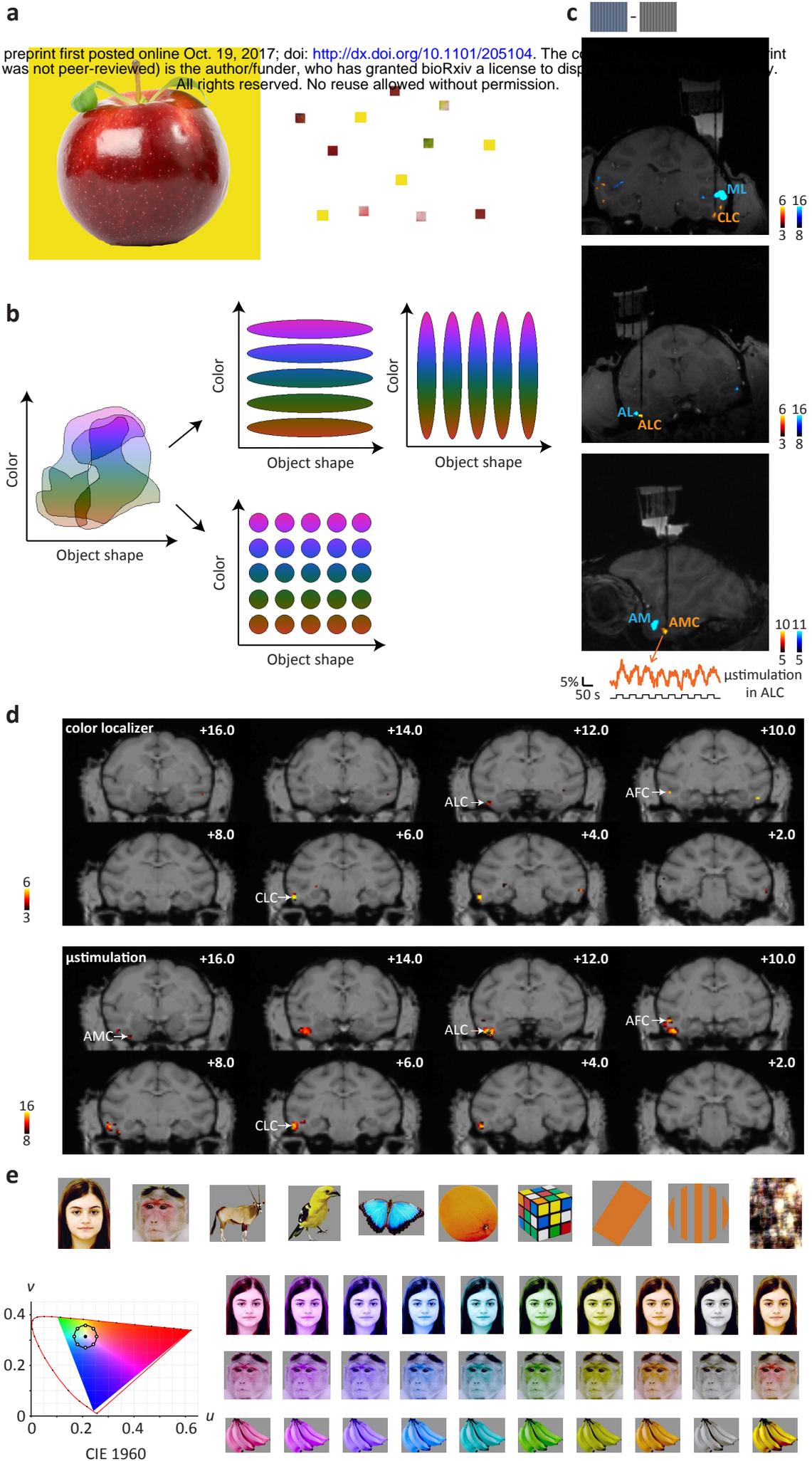
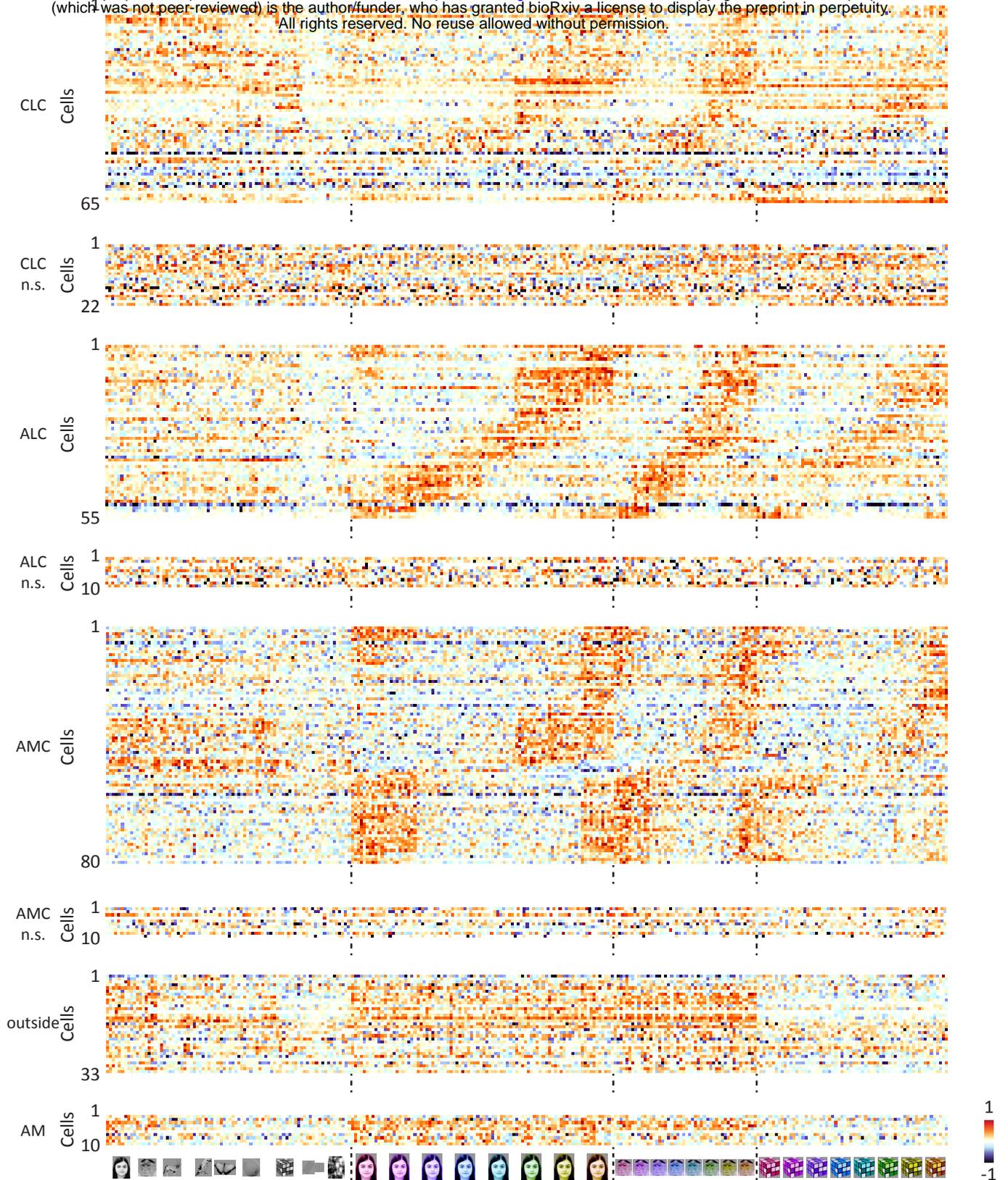
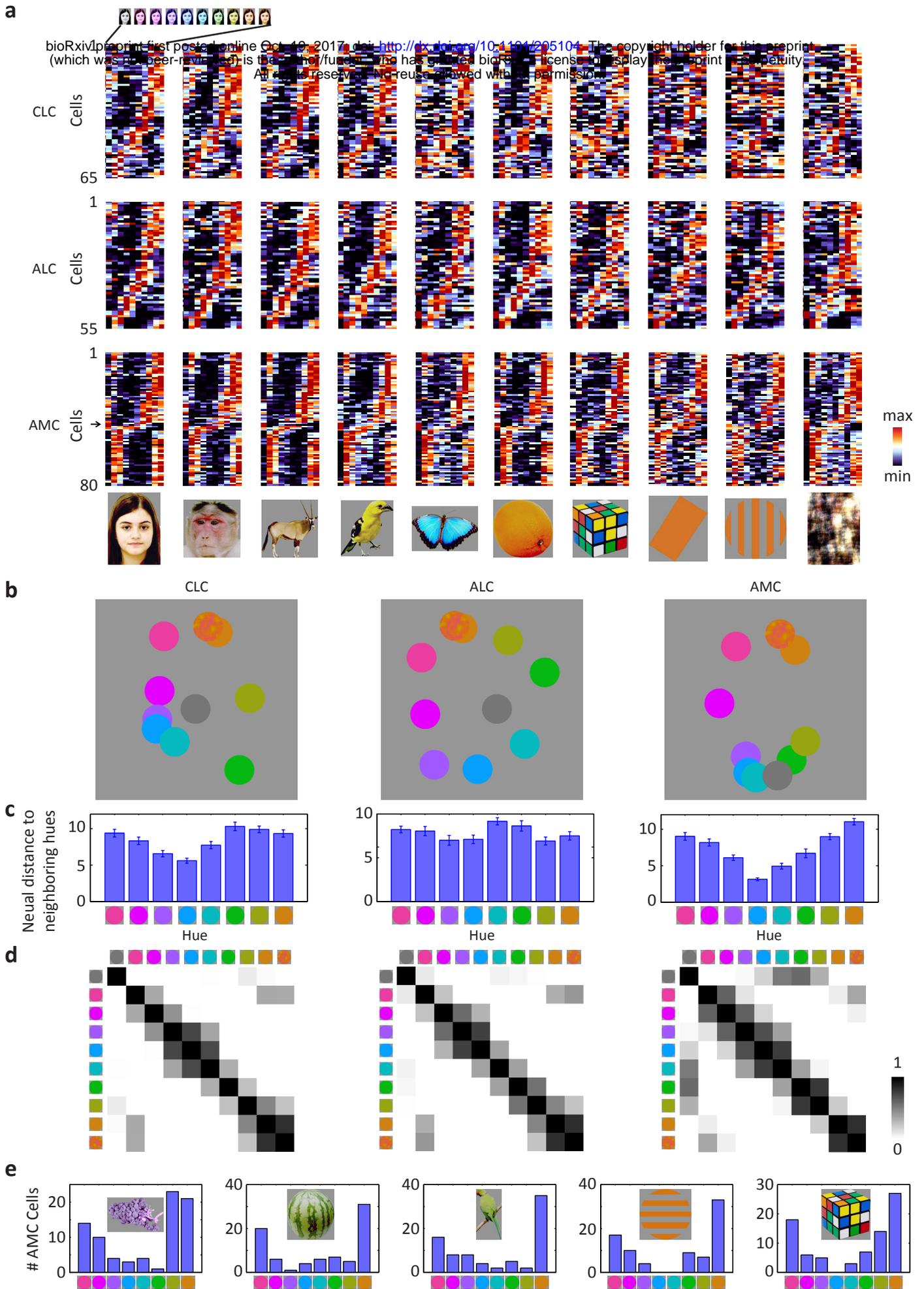
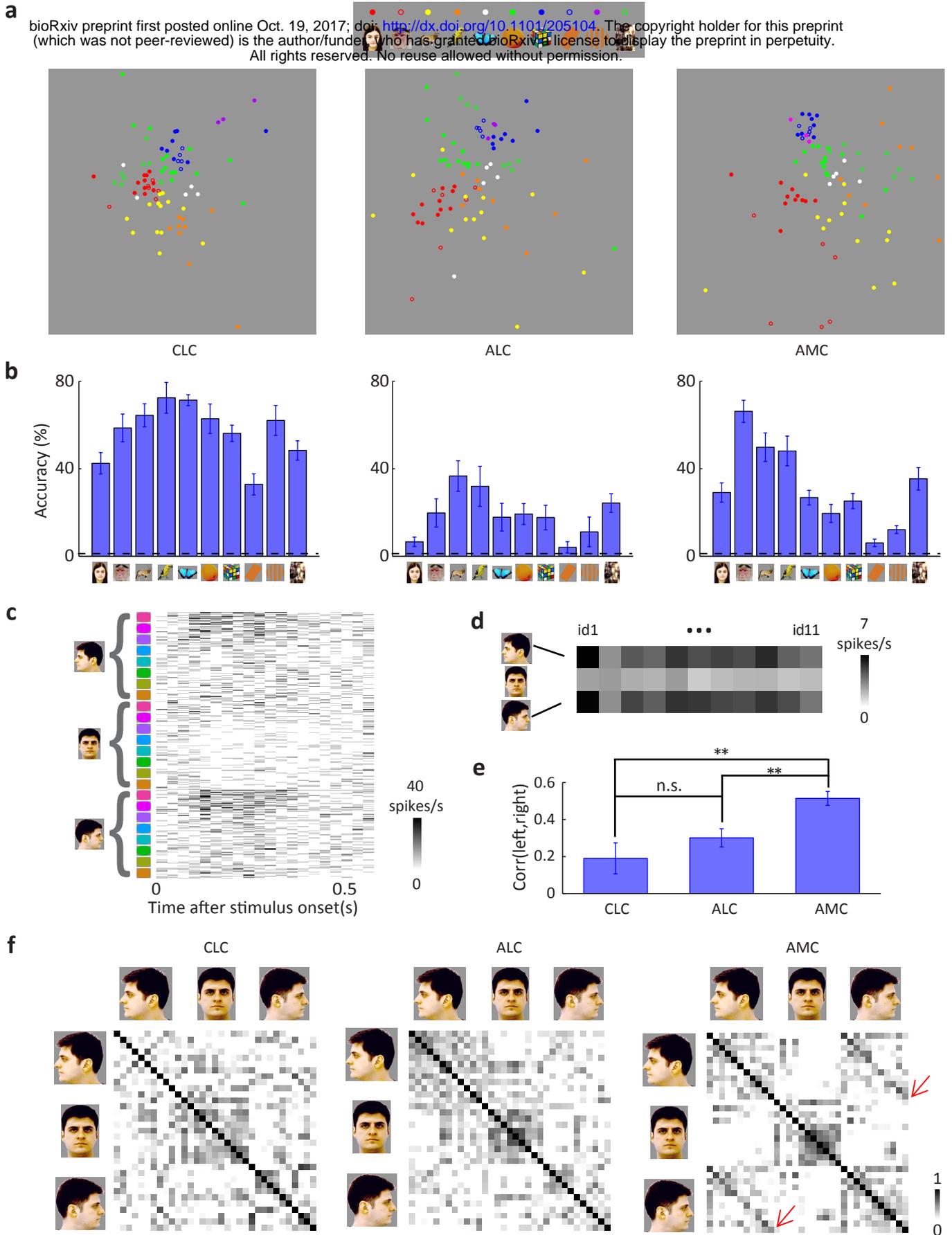


Figure 2

bioRxiv preprint first posted online Oct. 19, 2017; doi: <http://dx.doi.org/10.1101/205104>. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.







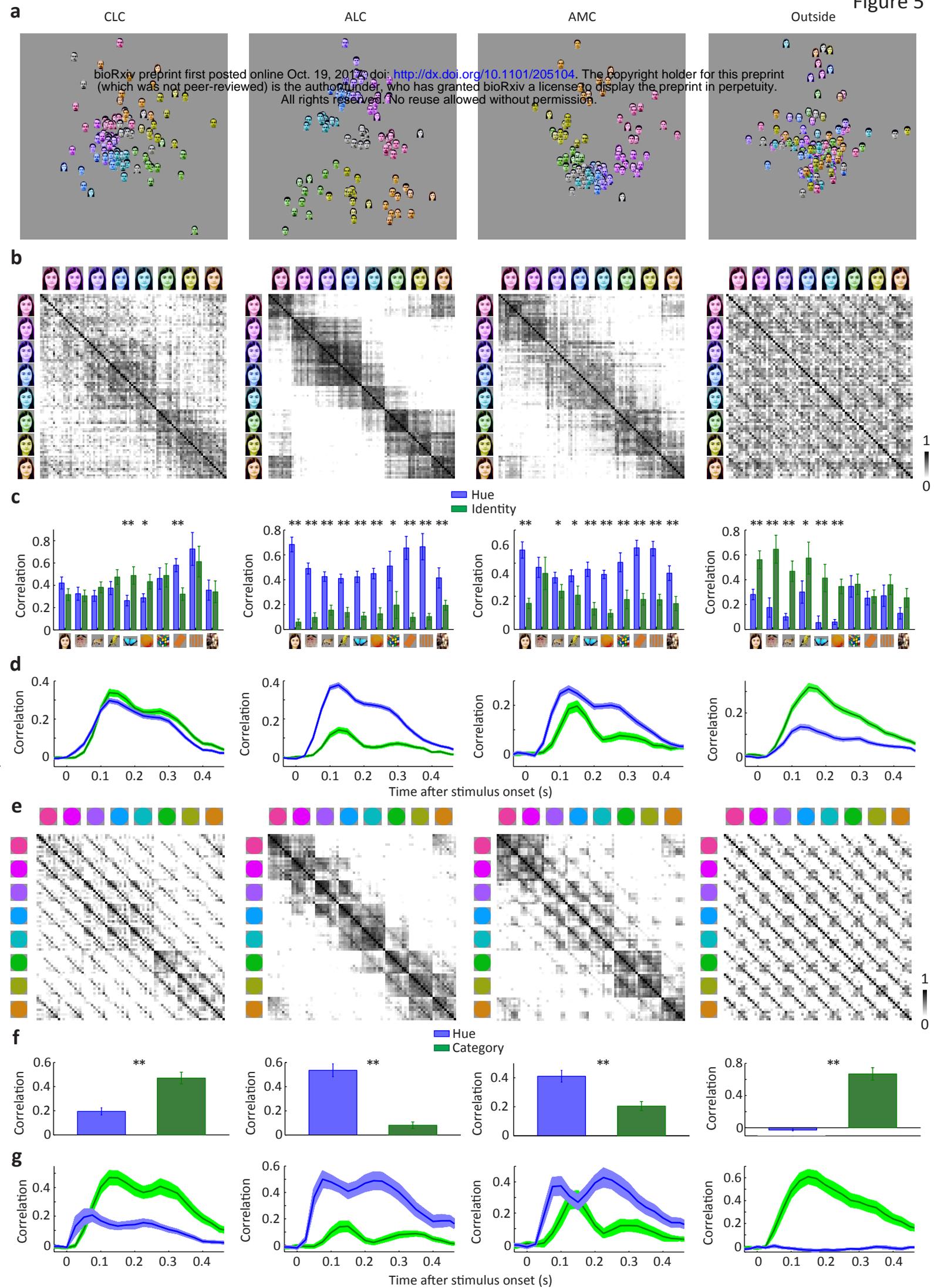


Figure 6

bioRxiv preprint first posted online Oct. 19, 2017; doi: <http://dx.doi.org/10.1101/205104>. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

