1 **Title:** IMAGE SEGMENTATION BASED ON RELATIVE MOTION AND

2 RELATIVE DISPARITY CUES IN TOPOGRAPHICALLY ORGANIZED AREAS OF

3 HUMAN VISUAL CORTEX

4

5 **Abbreviated title:** RELATIVE MOTION AND DISPARITY IN VISUAL CORTEX

6

7 Peter J. Kohler*[a], Benoit R. Cottereau[b,c] and Anthony M. Norcia[a]

8

9 [a]Department of Psychology, Stanford University

10 Jordan Hall, Building 420, 450 Serra Mall, Stanford, CA 94305

11

12 [b] Université de Toulouse, Centre de Recherche Cerveau et Cognition, Toulouse, France

13 [c] Centre National de la Recherche Scientifique, Toulouse Cedex, France

14

15

16

17 ***Corresponding Author:** pjkohler@stanford.edu

18 **Number of pages:** 30

19 **Number of figures:** 6

20 **Number of words**

21 **Abstract:** 199

22 **Introduction:** 663

23 **Results:** 1932

24 **Discussion:** 1825

25 **Methods:** 2449

26 **Conflict-of-interest:** The authors declare no competing financial interests.

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

RUNNING TITLE: Relative motion and disparity

28   The borders between objects and their backgrounds create discontinuities in image
29   feature maps that can be used to recover object shape. Here we used functional magnetic
30   resonance imaging (fMRI) to study the sensitivity of visual cortex to two of the most
31   important image segmentation cues: relative motion and relative disparity. Relative
32   motion and disparity cues were isolated using random-dot kinematograms and
33   stereograms, respectively. For motion-defined boundaries, we found a strong
34   retinotopically organized representation of a 2-degree radius motion-defined disk, starting
35   in V1 and extending though V2 and V3. In the surrounding region, we observed phase-
36   inverted activations indicative of suppression, extending out to at least 6 degrees of
37   retinal eccentricity. For relative disparity, figure responses were only robust in V3, while
38   suppression was observed in all early visual areas. When attention was captured at
39   fixation, figure responses persisted while suppression did not, suggesting that suppression
40   is generated by attentional feedback from higher-order visual areas. Outside of the early
41   visual areas, several areas were sensitive to both types of cues, most notably hV4, LO1
42   and V3B, making them additional candidate areas for motion- and disparity-cue
43   combination. The overall pattern of extra-striate activations is consistent with recent
44   three-stream models of cortical organization.

45
46
47
48
49
50
51
52
53
54
55
56
57
58

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

59 **Introduction**

60 The boundaries between objects and their background give rise to discontinuities in

61 multiple feature maps. Relative motion and relative disparity are two strongly related

62 parallax cues, that each contribute independently to image segmentation and the

63 perception of shape, and can be combined to disambiguate 3D object and scene

64 structure[1-3]. Sensitivity to image discontinuities created by relative motion has been

65 observed in both early and higher-level visual areas. Single unit responses to relative

66 motion information are robust in primate V1 and V2[4-9], as well as in MT[10-13] and IT[14],

67 although selectivity for the orientation of a motion-defined boundary may not arise until

68 V2[9,15]. In humans, functional magnetic resonance imaging (fMRI) studies[16,17] identified

69 an area in dorsomedial occipital cortex, originally termed the kinetic occipital area (KO),

70 that was activated by relative motion in texture-defined bars. However, other human

71 fMRI studies at that time found that this stimulus also produces activations in V1, V2, V3

72 and in the hMT+ complex that includes the homologue of macaque MT[18-20]. Later work

73 made more extensive measurements in topographically organized visual areas and used

74 fMRI adaptation to identify selectivity for the orientation of motion boundaries in areas

75 V3A, V3B, LO1, LO2 and V7[21], which partially overlapped with functionally-defined

76 area KO.

77 Studies in macaque suggest that image discontinuities generated by relative

78 disparity are not encoded before V2[22-24]. Sensitivity to disparity discontinuities has also

79 been found in V3[25], V4[26,27], IT[28] and, depending on the precise definition of relative

80 disparity, in MT[29]. In humans, the functional localization of relative disparity processing

81 is less well established. Several studies have compared fMRI BOLD responses to

82 displays with multiple disparity planes to responses produced by a single depth-plane,

83 and found dominant activations in V3A[30-32]. Because these studies contrasted non-zero

84 disparity and zero disparity displays, the effects could be driven by tuning for absolute or

85 relative disparity, or for both. This confound was addressed by Neri and colleagues[33],

86 who used fMRI adaptation to measure responses to both absolute and relative disparity

87 separately. They found that areas V4 and V8 (and to a lesser degree early visual areas V1,

88 V2 and V3) showed adaptation effects for both relative and absolute disparity, but that

89 adaptation was only present for absolute disparity in V3A, MT and V7. A more recent

90   study of cue combination reported reliable classification of depth defined by motion and

91   disparity in several visual areas, but used an experimental design that allowed

92   classification to be driven by tuning for absolute or relative signals, or for both[1].

93   Here we seek to integrate this past work by measuring fMRI BOLD responses to

94   simple figure-ground displays based on relative motion and relative disparity in 18

95   topographically organized visual areas. We use disk-annulus stimulus configurations that

96   are well-matched between motion and disparity, and an experimental design that controls

97   for the contribution of absolute signals to the measured responses. The use of a figure-

98   ground stimulus configuration is also expected to drive global shape processes and thus

99   activate object-sensitive cortical areas[34-36]. Finally, we conduct a control experiment

100   where attention is captured at fixation, to investigate the influence of top-down attention

101   on responses.

102

103   **Results**

104   *Eccentricity analysis in early visual cortex.*

105   We begin by describing results from early visual areas V1, V2 and V3, each of

106   which have strong retinotopic representations of the visual field. To visualize and

107   quantify the existence and retinotopic specificity of the representations of the motion and

108   disparity-defined figures on the cortical surface we used a retinotopy template[37] to define

109   eccentricity-based sub-ROIs and then used a fMRI localizer task to verify the accuracy of

110   the template. The goal of the sub-ROI analysis was to track the responses to central

111   (figure), boundary and outer (ground) regions of the stimulus across the surfaces of the

112   early visual areas. The localizer task consisted of the alternation across 12 sec block of

113   high contrast dynamic textures in the disk with high contrast dynamic textures in the

114   background (see Fig. 1A). We computed the signed amplitude values over the period of

115   block alternation, by applying a Fourier transform to the average time course (see

116   Methods). The results of the contrast-based localizer condition are plotted in Figure 2A.

117   Positive amplitudes indicate relatively stronger responses to the A block, during which

118   the disc was presented, compared to the B block, during which the annulus was

119   presented. Negative amplitudes indicate the opposite. For convenience, we will refer to

120   the former as "disk responses" and the latter as "annulus responses". The eccentricity

121   axis was derived independently, based on the template.
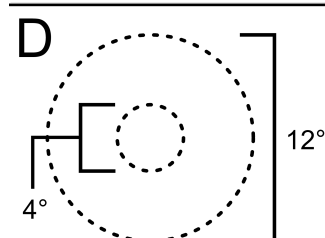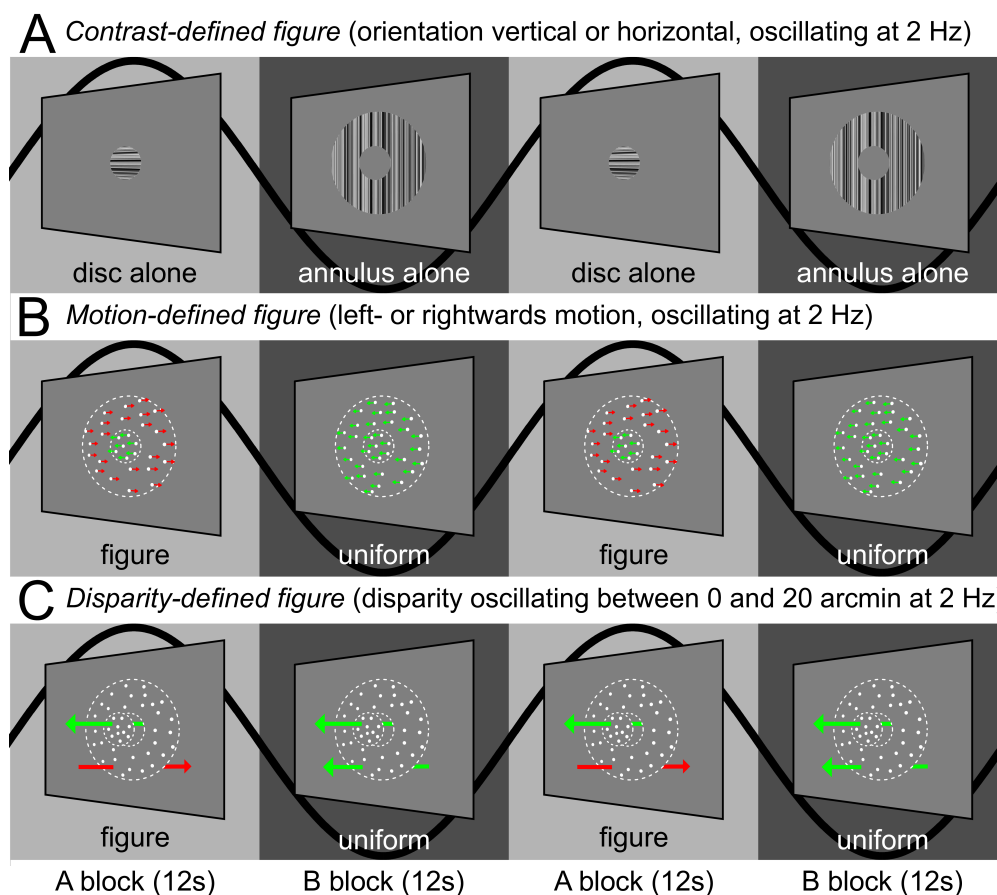

122
123
124   **Figure 1. The experimental design used in Experiment 1.** The on-off responses evoked by the conditions
125   that are captured by our Fourier analysis, are illustrated with sine waves. (A) The contrast condition – the
126   orientation of both the disc and annulus alternated between vertical and horizontal at 2 Hz. Disk and
127   annulus were presented in temporal succession. (B) The motion condition – dots alternated between left-
128   and right-wards motion at 2Hz. When the disc and annulus regions alternated in anti-phase, a figure percept
129   was evoked, when alternations were in-phase, there was no such percept. (C) The disparity condition – the
130   dots alternated between the fixation plane (0 arcmin) and a position behind fixation (uncrossed disparity, 20
131   arcmin) at 2 Hz. Anti-phase movement of the disc and annulus regions lead to a figure percept, while in-
132   phase movement generated no such percept. The white dotted line in B and C indicates the extent of the
133   disc and annulus used in A. Note that the relative disparity condition was identical to that used in
134   Experiment 2, except for the introduction of an RSVP task at fixation, which directed attention away from
135   the stimulus. (D) Schematic outlining the size of the disc and annulus region, respectively, which was
136   shared across all conditions.
137

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

138   In areas of cortex defined by the template as responsive to eccentricities covered

139   by the disc, we see a high amplitude disc responses. At eccentricities defined as near the

140   disc boundary (2º radius), the phase sign reverses and we see annulus responses. There

141   are significant responses at all eccentricities covered by the stimulus, except for the sign

142   reversal region, where voxels responding to the disc and the annulus, 180º out of phase,

143   are likely mixed together. The eccentricity of the sign reversal is slightly biased towards

144   the periphery (about the size of the separation between sub-ROIs, 0.25º), but otherwise

145   the pattern of results indicates that the template-based procedure allows us to accurately

146   localize the eccentricity of the contrast-defined boundary in retinotopic cortex. The 180º

147   phase reversal can also be directly observed in the whole-brain analysis of the contrast

148   condition (Figure 6A), which plots the vector-averaged phase across all participants.

149   While the contrast condition alternated a disc and an annulus, the motion and

150   disparity conditions alternated figure-ground and a uniform field configurations. The dots

151   updated dynamically at 20 Hz throughout each block. In the disparity condition, they

152   were temporally uncorrelated across updates, and thus did not generate monocular cues to

153   form. In the motion case, the dots were temporally correlated with a long lifetime. The

154   global structure of the displays (uniform vs segmented) updated at 2 Hz, alternating

155   between left- and rightwards movement in the motion condition, and between the fixation

156   plane (0 arcmin), and a position behind fixation (uncrossed disparity, 20 arcmin) in the

157   disparity condition. The disc and annulus region alternated in anti-phase during the A

158   block, and in phase during the B block. Importantly, absolute motion and absolute

159   disparity were updating at 2 Hz during both A and B blocks, but only the A block

160   generated relative motion and disparity cues that give rise to a figure boundary. The

161   design and corresponding Fourier analysis yields the differential response and thus allows

162   us to isolate responses driven by the presence of relative motion/disparity discontinuity,

163   the segmented figure surface or some combination of the two from absolute

164   motion/disparity responses.

165   The overall response amplitudes in the motion and disparity conditions were 5-10

166   times weaker than in the contrast condition, but there were nonetheless significant

167   activations at multiple eccentricities. For both motion and disparity, positive amplitudes

168   indicate stronger responses to the A block in which the figure and surround were

169   segmented ("figure responses"), while negative amplitudes indicate stronger responses to

170   the B block in which a uniform field was perceived ("uniform responses"). For motion

171   (see Figure 2B), we saw figure responses at the disc-annulus boundary, which persisted

172   beyond the extent of the disc by 1º in V1 and V2, and even further in V3. There were also

173   figure responses inside the disc region, which were most prominent in V2 and V3. This

174   result indicates that we can detect responses to a figure boundary defined by relative

175   motion in all three early visual areas. Surprisingly, we also saw uniform responses at

176   eccentricities > 4.5º, which were significant in all visual areas, although less consistently

177   in V1. This results indicates that when the boundary is present, responses in the region

178   outside the boundary are weaker than when the boundary is not present. This is likely due

179   to suppression of the surround when the boundary is present, rather than surround-

180   exclusive enhancement of responses when the field is uniform. The lack of consistent

181   suppression in V1 may be due to insufficient sensitivity of our method.
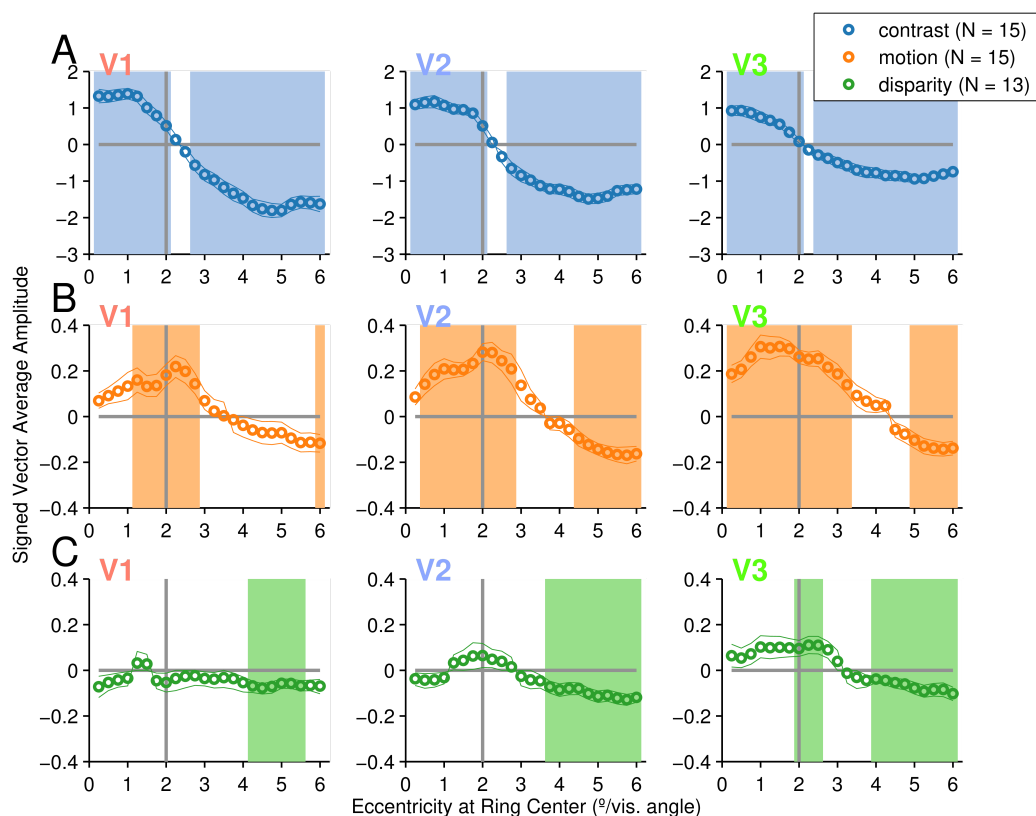

182

183   **Figure 2: Eccentricity analysis of V1-V3 in Experiment 1.** Signed vector mean amplitudes for the
184   contrast (A), motion (B) and disparity (C) conditions within 24 eccentricity-defined sub-ROIs, centered on
185   eccentricities spaced 0.25º apart, each having a width of 0.5º. The shaded areas on the plots indicate
186   condition ✕ sub-ROI combinations that were significant at $\alpha = 0.05$. The color of the text for the ROI
187   names matches the ROI colors on Figures 4 and 6.

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

188   In the disparity condition (see Figure 2C), we did not measure significant figure

189   responses in V1 or V2 – although V2 responses at the boundary were positively signed,

190   they were clearly not significant (lowest $p = 0.445$). In V3 by contrast, there were

191   significant figure responses starting at the boundary and extending to ~0.5º outside the

192   boundary. We also saw significant uniform responses in all early visual areas, which, as

193   for the motion condition, are consistent with suppression that occurs during the A block

194   when the boundary is present, but not in the B block when the field is uniform.

195   These results suggest an interesting dichotomy: Among the early visual areas, we

196   only see evidence for responses to a figure boundary defined by relative disparity in V3,

197   while all three early visual areas have evidence of suppression when the disparity-defined

198   boundary is present. Relative suppression of surround-region responses may be due to

199   feedback from higher cortical areas, driven by top-down attentional selection. Spatial

200   attention can modulate BOLD responses in all early visual areas, including V1[38-40]. In

201   Experiment 2, we directly tested the hypothesis that the suppression is due to attention-

202   driven feedback.

203   Experiment 2 used the same parameters as the relative disparity condition in

204   Experiment 1, except that attention was now directed away from the stimulus via an

205   orthogonal task presented at fixation (see the 'Visual stimuli' section in the Methods).

206   Under these conditions, we measured significant figure responses in V3, but not in V1

207   and V2, replicating results from Experiment 1 (see Figure 3). A single uniform response

208   reached significance at the inner-most eccentricity in V2, but given the lack of spatial

209   correspondence with the stimulus boundary and absence of any other trends in the V2

210   data, we consider this type 1 error. Importantly, we saw no evidence of a uniform

211   response indicating suppression anywhere in any of the early visual cortex ROIs. This

212   result support our hypothesis that the uniform response we saw in the surround region in

213   Experiment 1 was due to top-down attentional suppression of the surround. Note that the

214   figure-region response extended even further beyond the figure (~1.5º) than it did in

215   Experiment 1 (~0.5º), perhaps due to a reduction of the negatively signed uniform

216   responses in the surround. Our results also support the conclusion that the sensitivity of

217   area V3 to a relative disparity-defined boundary that we measure as positively signed

218   responses is not strongly dependent on attention.

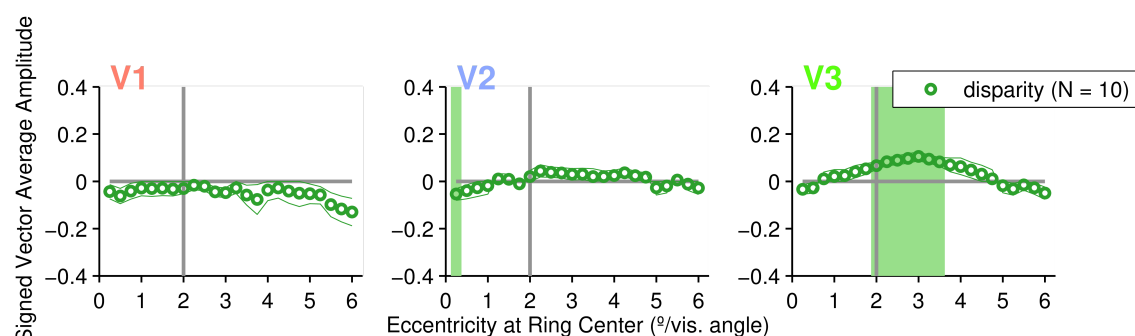THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

219

Figure 3: Eccentricity analysis of V1-V3 in Experiment 2. Signed vector mean amplitudes within 24 eccentricity-defined sub-ROIs, centered on eccentricities spaced 0.25° apart, each having a width of 0.5°. The shaded areas on the plots indicate condition $\times$ sub-ROI combinations that were significant at $\alpha = 0.05$. The color of the text for the ROI names matches the ROI colors on Figures 4 and 6. The relative disparity condition shown here was identical to the one used in Experiment 1 (see Figure 2C), except attention was directed away from the stimulus.

Extended ROI-based analysis of responsivity

The next set of analyses quantifies motion and disparity sensitivity in topographically organized visual areas beyond early visual cortex. Activation was defined as the signed vector mean amplitude of the average time course across all voxels within each ROI. The logic was the same as for the sub-ROI within early visual cortex: Positive amplitudes indicate responses to the figure (A block), while negative amplitudes indicate stronger responses to the uniform field (B block).

In both the motion- and disparity-defined form conditions, all significant activations were positively signed. This is consistent either with positive-sign activations generated by the figure overcoming any negative-sign activations that may have occurred in a subset of voxels within a given ROI, or a lack of negative-sign activation. We distinguished three response patterns: areas that only had significant responses ($p < 0.05$; indicated with shaded areas in Figure 4) to the motion-defined figure (hV4, TO1), areas that only had significant responses to the disparity-defined figure (PHC1, PHC2, IPS1, IPS2) and areas that had significant responses to both (V3B, LO1, LO2, IPS0, IPS3). Our results suggest a clear functional distinction between V3B and nearby area V3A, which has no significant responses. Motion generated stronger responses than disparity in V3 (see Figure 2), but this was less pronounced among the higher-level areas: Both V3B, IPS0 and IPS3 had nearly identical responses to the two conditions.

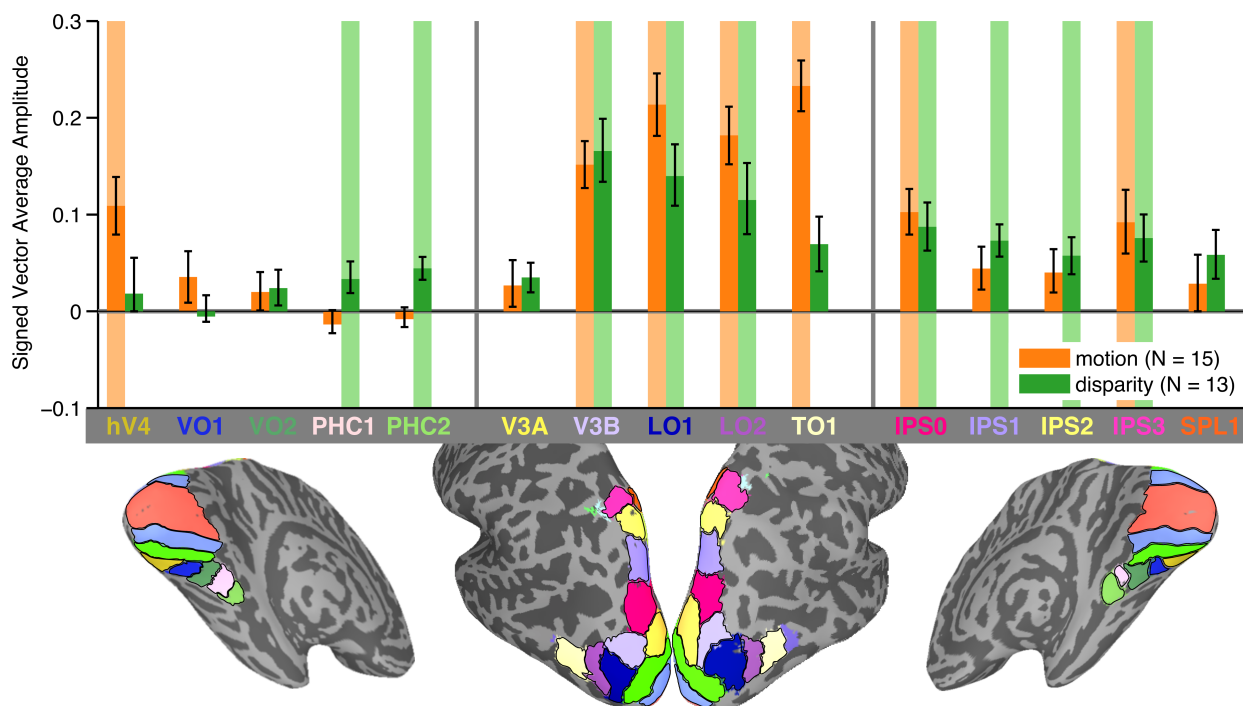THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

Figure 4: ROI results from Experiment 1. Signed vector mean amplitude within 15 topographically organized regions of interest, excluding early visual areas V1-V3, which are shown in Figure 2. The motion condition is shown in orange, and the disparity condition in green. Shaded areas behind the bars indicate condition × ROI combinations that were significant at α = 0.05. The ROIs are shown for the both hemispheres of an example participant's inflated cortical surface reconstruction, below the graph.

Effect of attention on disparity-defined figure activations outside of early visual cortex

We now ask if the areas that were sensitive to a disparity-defined figure boundary in Experiment 1, were also significant in Experiment 2, when attention was directed away from the stimulus via an orthogonal task at fixation. We found that responses persisted in areas V3B and LO1, but not in LO2 and IPS0-3 (see Figure 5), suggesting that activations in these latter areas depend on attention. Note that we were unable to probe PCH1 and 2 as these ROIs that were not covered by the fMRI acquisition protocol used in Experiment 2. hV4 also had a significant response ($p = 0.008$), which we did not see in Experiment 1 ($p = 0.893$). This difference in activation patterns could occur if hV4 has negative-sign, attention-dependent surround activations that cancel out the positive-sign figure activations. Our eccentricity analysis of early visual cortex demonstrated that these negatively-signed attention-driven effects are eliminated when top-down attention is controlled in Experiment 2. It is likely that the same thing is happening in hV4, eliminating the cancellation and revealing the positive-sign activations.

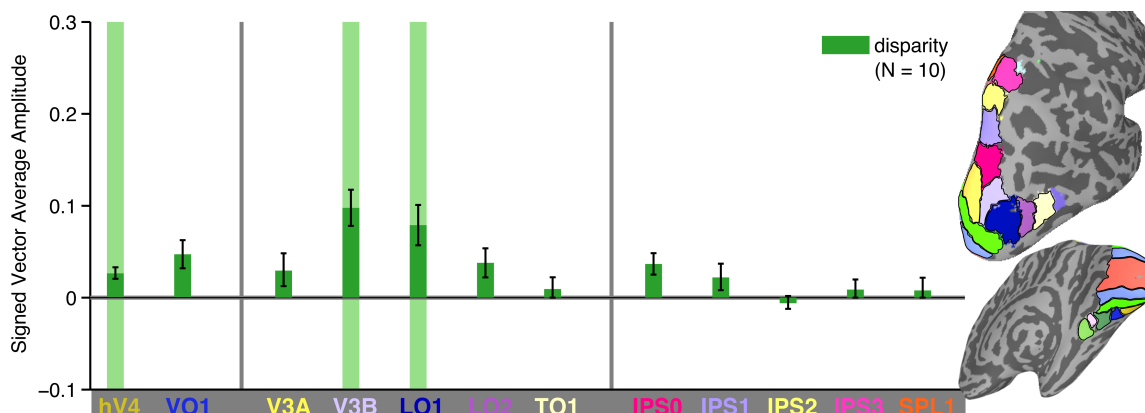THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

Figure 5: ROI results from Experiment 2. Signed vector mean amplitude within 12 topographically organized regions of interest, excluding early visual areas V1-V3, which are shown in Figure 3, and three ROIs that were not covered by the fMRI acquisition protocol used in Experiment 2 (VO2, PHC1 and PHC2). Shared areas behind the bars indicate condition × ROI combinations that were significant at α = 0.05. The ROIs are shown for the right hemisphere of an example participant's inflated cortical surface reconstruction, on the right side of the graph.

Whole brain analysis

A surface-based alignment approach was used to visualize vector-averaged responses to the conditions in Experiment 1, across all of cortex, including regions outside our set of ROIs. The results of these analyses were largely consistent with the sub-ROI and whole-ROI analyses, and we will only describe them briefly.

We plot the phase of the vector-averaged response, thresholded by significance. Blue colors indicate responses in phase with the A block, while orange colors indicate responses to the B block. For the contrast condition (see Figure 6A), the reversal of the phase sign from disc responses to annulus responses between low and high eccentricities, described in the sub-ROI analysis (see Figure 2A) can be clearly observed in early visual areas.

For motion, we see clear evidence in early visual cortex of both the figure region response at low eccentricities (Figure 6B, blue colors) and uniform responses consistent with the suppression of the surround (orange colors) that was described in the sub-ROI analysis. We also see figure responses that cover most dorsolateral ROIs and extend anteriorly and ventrally beyond the ROIs. In correspondence with the whole-ROI analysis (see Figure 4), we also see figure responses in IPS0, IPS3, and hV4, but not in V3A.

294 **Figure 6: Whole-brain results from Experiment 1.**
295 Vector-averaged phase maps, thresholded at $\alpha = 0.05$.
296 Maps were produced through surface-based alignment
297 procedure in which each subjects' cortical mesh was
298 converted to a standardized mesh, which allowed for cross-
299 subject comparisons of values at each mesh node. The
300 vector average phase across subjects, as well as a
301 corresponding p-value based on both amplitude and phase,
302 could then be computed for each mesh node. Blue shades
303 indicate responses to the A block, while orange shades
304 indicate responses to the B block. The ROIs are outlined
305 on the surface, and labeled version can be inspected in
306 Figure 4 and 5.
307

308   For disparity, we again see uniform
309 responses consistent with suppression in early
310 visual cortex (see Figure 6C, orange colors). It is
311 worth noting that for both motion and disparity,
312 there is little evidence of suppression outside of
313 early visual cortex, at least within our ROIs. We
314 do not see evidence of figure responses in V3,
315 likely because the surface-based alignment
316 approach is less sensitive than the sub-ROI
317 analysis (see Figure 2). We do see figure
318 responses in V3B, LO1, and LO2 and several
319 IPS areas, but not in TO1 and V3A, in
320 correspondence with the whole-ROI analysis (see
321 Figure 4).
322

4. Discussion

324 We find evidence for representations of relative
325 motion-defined figure boundaries in all early
326 visual areas, and evidence for an analogous
327 representation defined by relative disparity at least as early as V3, largely consistent with
328 prior work in macaque[22,23]. In these areas, the activation patterns for both cues reflect the
329 visual field topography of the stimulus, including a region of enhanced responses at or
330 near the figure boundary, surrounded by suppressed responses in the ground region. We
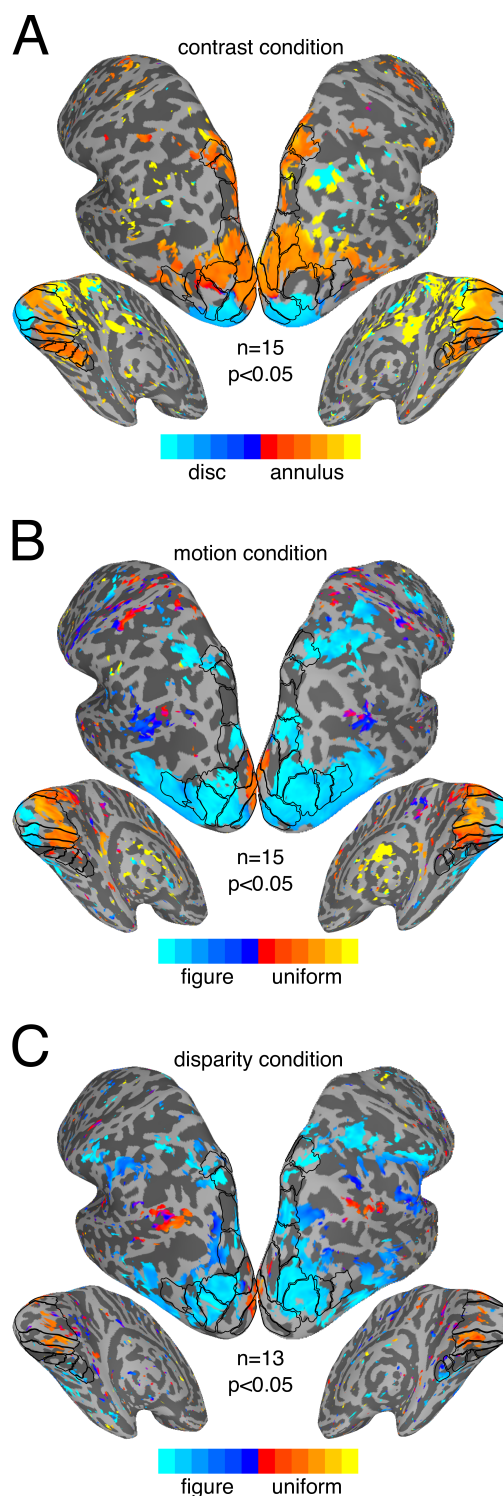
331   measure significant suppression associated with relative disparity in all early visual areas,

332   but enhancement only reaches significance in V3, which may be due to lack of sensitivity

333   of our method.

334   This pattern of results suggests that early visual areas go beyond simple edge-detection as

335   reported previously in macaque V2[23], and maintain representations related to the

336   perceptual organization of the stimulus into figure and ground regions, as has been

337   suggested for disparity on the basis of other single-unit recordings in macaque V2[24].

338   Outside of early visual areas, we found areas that are sensitive to relative motion

339   (TO1), relative disparity (PHC1, PHC2, IPS1, IPS2) or both (V3, hV4, LO1, LO2, V3B,

340   IPS0, IPS3). Areas that are jointly selective contain candidate substrates for integrating

341   these two cues. Ban and colleagues[1] found evidence of integration of motion and

342   disparity in V3B/KO, but not in any of the other five areas that we find to be sensitive to

343   both cues. In Experiment 2, when attention was directed away from the stimulus using a

344   letter task at fixation, disparity tuning was eliminated in a subset of the areas (LO2, IPS0-

345   3), indicating that in those areas, sensitivity to disparity depends on attention. In V3, hV4,

346   V3B and LO1, disparity tuning was independent of attention.

347   We observed stronger responses to relative motion compared to relative disparity

348   in early visual areas (see Figure 2), but this difference was less pronounced in higher-

349   level areas. This result is consistent with the 'single-cue' classification accuracies

350   reported by Ban and colleagues, which are greater for motion than disparity in early

351   visual areas, but more comparable in higher-level areas[1]. In some areas, their 'single cue'

352   classification accuracies are above-chance for stimuli that do not produce significant

353   responses in our data, most prominently in V1, V2 (disparity) and V3A (motion and

354   disparity). It may be that in those cases, the classifier is picking up tuning for absolute

355   motion and disparity[1], which was controlled in our experiment design. Thus, the pattern

356   of results in our current study and the study by Ban and colleagues[1] suggest that V3A is

357   sensitive to absolute, but not relative, motion and disparity, while neighboring V3B is

358   sensitive to relative motion and disparity.

359   Relative disparity processing has been associated with the "canonical macaque

360   ventral stream" leading from V4 to IT[41-43]. We find hV4 to be responsive to relative

361   disparity as well as relative motion cues, which is consistent with reports that macaque

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

362   V4 is sensitive to both relative disparity[26,27,44], and relative motion[45]. In the macaque

363   ventral stream, areas down-stream of V4 include sub-divisions of IT, TEO and TE, that

364   are responsive to relative disparity[46,47] and to relative motion[14,48]. In our data, areas

365   immediately downstream of hV4, VO1 and VO2, are responsive to neither cue, while

366   lateral surface areas LO1 and LO2 are responsive to both. Thus, for motion- and

367   disparity-defined contours, the functional homology is poor between human ventral

368   surface areas and macaque ventral areas, but better between human lateral surface areas

369   and macaque ventral areas. Two recent proposals divide human visual cortex into three

370   (dorsal, ventral and lateral) rather than the two, canonical dorsal and ventral streams of

371   the macaque[49,50]. Our data suggests that functional homology exists between human

372   lateral areas and the macaque ventral areas in IT cortex. Human ventral surface areas may

373   either lack clear homologues in the macaque, or the homologous areas remain to be

374   discovered.

375   Activations in V1, V2 and V3 for motion- and disparity-defined figures each

376   consisted of a region of positive-sign activation surrounded by a negative-sign activation.

377   The negative-sign activation in the surround could arise from suppression of responses

378   during the A block or enhanced responses during the B block. There are, however,

379   several reasons to suspect that suppression is the main driver of our negative-sign

380   activation. First, the center and surround region are indistinguishable in the uniform

381   stimuli used in the B blocks, and it is thus unlikely that there would be exclusive

382   enhancement of the surround that would manifest as a negative-sign activation. Second,

383   data from both macaque electrophysiology and human fMRI suggests that the relatively

384   negative activations we see in the ground-region reflect suppression of neural responses.

385   An opposing sign organization has been observed previously for "first-order" contrast-

386   defined stimuli by contrasting blank screen baseline activation level to that within and

387   adjacent to the retinotopic location of high contrast patterns[51-53]. Single unit and local

388   field potential recordings in the retinotopic and adjacent representations of spatially

389   localized, high contrast stimuli indicate suppression of baseline neural responses in the

390   regions that had negatively signed BOLD[53]. Similar suppression may be occurring with

391   our "second-order" contrast stimuli.

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

392   The center-surround configuration we used lends itself to the detection of

393 alternate-sign activations because eccentricity is mapped systematically on the cortical

394 surface in the foveal confluence-region of early visual cortex[54]. A suggestion of negative-

395 sign or "out-of-phase" BOLD activation was present in the disparity data of Parker and

396 Bridge[55], but their use of rotating wedge-shaped stimuli complicated its visualization and

397 measurement. Negative-sign activation has also been observed with a display in which a

398 second-order figure region (a bar) was defined by temporal transients[56]. Negative-sign

399 activation was found adjacent to the retinotopic locus of the bar, but that study did not

400 find enhanced activation within the retinotopic representation of the figure as indicated

401 by the positive-sign activations we observed[57]. Reppas and colleagues[18] observed

402 positive-sign activation at the border of a motion-defined form, but found neither

403 negative-sign activation in the ground region nor positive-sign activation within the

404 figure region, as we see in our data. Next, we will consider how suppression in the

405 ground region could be driven by feedback from higher visual areas.

406   In Experiment 2, participants performed a demanding task at fixation that diverted

407 attention away from the changing disparity stimuli. Under these conditions, the negative-

408 sign activation in the surround was unmeasurable in all three visual areas, while the

409 positive-sign activation at the figure region in V3 persisted. This suggests that the

410 suppression was driven by feedback related to top-down attention, which decreased when

411 participants no longer attended the figure. BOLD responses can be modulated by

412 attention in a spatially specific fashion as early as V1[38-40]. If the surround suppression we

413 observe is in fact due to attention-related feedback, we would expect to see it throughout

414 early visual cortex, even in areas that are not sensitive to a boundary defined by relative

415 disparity. The fact that negative-sign activation is observable for relative motion stimuli

416 even at 6 degrees of retinal eccentricity in V1 is consistent with feedback from higher-

417 level visual areas, as population receptive fields at this eccentricity, measured using

418 fMRI, are on the order of 2 degrees or less[58,59]. Finally, we note that disparity tuning in

419 IPS areas also disappeared when attention was captured at fixation, suggesting that these

420 areas may be sources of feedback to early visual areas.

421   A surprising aspect of our data is the apparent expansion of the enhanced

422 responses associated with the figure-region for disparity and motion-defined figures.

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

423   Here we propose an explanation for these results, based on the way that fMRI voxels

424   sample the visual field. Population receptive fields (pRFs) reflect a summary of the

425   receptive fields of neurons sampled by each fMRI voxel,[60] and are already quite large at

426   2º (>0.5º in V1, larger in V2 and V3[59]). This means that sub-ROIs both inside and outside

427   the boundary will contain a mixture of voxels with pRFs that overlap with the boundary,

428   and voxels with pRFs that do not. In the contrast condition, when the B block drives the

429   surround as much as the A block drives the center, we would expect the sign reversal to

430   occur at the boundary. It in fact occurs just outside (see Figure 2A), a bias that is likely

431   due to the fact pRFs increase in size towards the periphery. This means that it is more

432   likely that pRFs centered outside the disc will overlap with the disc than it is that pRFs

433   centered inside the disc will overlap with the annulus, resulting in a bias in disc responses

434   towards the periphery.

435   For the motion and disparity conditions in Experiment 1, the A block is driving

436   the boundary while the surround is suppressed, leading to a phase sign reversal. If the

437   relative suppression of the surround is weaker than the relative enhancement of the

438   boundary, the enhancement might overpower the suppression, both in the summation of

439   responses making up each pRF and in the averaging of pRFs within a sub-ROI. This

440   would lead significant boundary enhancement to 'survive' outside the boundary; which is

441   what we see in the motion and disparity conditions (see Figure 2B and C). If the B block

442   response is further weakened, the reversal should move further towards the periphery,

443   which is what we see in the disparity condition of Experiment 2 (see Figure 3). This

444   simple model explains the boundary responses observed outside the figure and cannot be

445   ruled out by our data. It is worth noting that the responses inside the figure, observed

446   most prominently for motion in V2 and V3, properly reflect true enhancement of the

447   figure, rather than pRFs sampling the boundary. pRFs are smaller towards the fovea and

448   therefore less likely to overlap with the boundary.

449   Figure-ground segmentation could be supported by several mechanisms acting

450   separately or in concert, including, as noted by Likova and Tyler[56], retinotopic

451   enhancement of the borders of the figure, retinotopic enhancement of the figure surface,

452   retinotopic suppression of the ground region, and a combination of figure enhancement

453   and ground suppression. There is both psychophysical[61] and fMRI[57] evidence for

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

454   "competition-mediated ground suppression" in the absence of figure enhancement. A

455   similar pattern of isolated suppression in the ground region was found with temporally

456   defined forms[56]. Evidence for spatially antagonistic facilitative/suppressive interactions

457   in figure-ground segmentation comes from an EEG source-imaging study[36] in which

458   relative disparity displays analogous to the one in the current study (segmented vs

459   uniform) were contrasted with displays in which the center moved in depth within an

460   uncorrelated surround and the complement (surround moving in depth, center

461   uncorrelated). The response patterns could be modeled by a multiplicative interaction

462   between center and surround in which the center response was enhanced and the surround

463   response was suppressed in proportion to the magnitude of relative disparity. Our data

464   from disparity- and motion-defined figures, combined with the single-unit

465   electrophysiology for texture-defined form[62] suggests that spatially antagonistic

466   interactions may be a general computational strategy used across multiple stimulus

467   domains. Our results are also consistent with the suggestion that boundary detection

468   circuits in early visual cortex provide a structure for attentional selection[63].

469   **Methods**

470   *Participants*

471   15 healthy adult participants (5 female; mean age = 30.6±13.5) participated in

472   Experiment 1 and 10 participated in Experiment 2 (3 female; mean age = 37.4±15.3),

473   with 1 participant taking part in both experiments. Each participant had visual acuity that

474   was better than +0.1 LogMar (20/25) in each eye as measured on a Bailey-Lovie chart

475   and stereo-acuity of 40 seconds of arc or better on the RandDot stereoacuity test. The

476   experiment began after the procedures of the study had been explained and the participant

477   had given written informed consent. Experiment protocol and consent forms were

478   approved by the Stanford University Institutional Review Board, and all methods were

479   performed in accordance with the relevant guidelines and regulations.

480

481   *Visual Stimuli*

482   The stimuli for Experiment 1 were shown on a 47" Resonance Technology LCD

483   display and viewed through a mirror at a distance of 277 cm. This resulted in a

484   presentation area of 12.1 × 21.2 °/visual angle, of which our stimuli occupied 12 × 12°.

485   The screen resolution was 1024 × 768 pixels, 8-bit color depth and a refresh rate of 60

486   Hz. In the relative disparity condition of Experiment 1, the mean luminance was 2.17

487   cd/m^2 and contrast was 60%, and stereoscopic stimuli were displayed using red/blue

488   anaglyph glasses, which were worn throughout the experiment. In the other two

489   conditions of Experiment 1, the mean luminance was 34.49 cd/m^2 and contrast was 90%.

490   In Experiment 2, the stimulus was viewed through Resonance Technology LCD goggles,

491   but the display parameters otherwise matched those used in the relative disparity

492   condition of Experiment 1.

493   For each stimulus condition, the display comprised a central 2° radius disk region

494   and an immediately adjacent 6° radius annulus. In the relative motion condition, the disk-

495   boundary was defined using a random-dot kinematogram. In the relative disparity

496   condition, the disk and the annulus were defined using dynamic random-dot stereograms

497   with no monocular cues. For the relative disparity display, dot size was 5 minutes of arc

498   (arcmin) and dot density was 36 per (°/visual angle)^2, while in the relative motion display,

499   dot size was 10.4 arcmin and dot density was 10 per (°/visual angle)^2. We also ran a

500   boundary localizer condition in which the disk-annulus boundary was defined by a

501   contrast difference in texture patterns comprised of 1-dimensional noise which alternated

502   between horizontal and vertical orientations at 3 Hz. This condition allowed us to

503   compare the boundary activations found in the motion and disparity conditions to the

504   activations generated by a contrast-defined boundary to which all visual areas should be

505   highly sensitive. This localizer also served to verify the accuracy of the retinotopy

506   template[37] we used.

507   In the contrast condition, the central disk was presented in what we will refer to as

508   the "A block" of the fMRI design and alternated with the adjacent annulus configuration,

509   presented during the "B block" (see Figure 1A). In the relative motion condition, the

510   horizontal positions of individual dots comprising a random dot pattern updated at 20 Hz.

511   Dots were displaced by 10 arcmin per update (3.33 °/sec) with a dot life-time of 100

512   video frames. The dots moved leftwards or rightwards, changing direction at 1 Hz. The

513   direction of motion inside and outside the central disk could either be in anti-phase,

514   leading to a spatially segmented percept with a visible boundary between the disc and

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

515 annulus regions defined by relative motion, or in phase, leading to a uniform motion

516 percept with no boundary. In the A block of the fMRI design, the display alternated in a

517 square-wave fashion between uniform motion and segmented configurations at 1 Hz. In

518 the B block, only uniform motion was shown. Locally, each part of the display contained

519 dots that alternated between leftward and rightward motion, only the relative direction of

520 motion over the disk and annulus regions differed between A and B blocks.

521 In the relative disparity condition, the positions of individual dots updated at 20

522 Hz such that the dot fields were binocularly correlated but temporally uncorrelated (no

523 monocular cues). The horizontal disparity of the central disk and the annulus alternated at

524 2 Hz between 0 disparity and 20 arcmin of uncrossed disparity. In the A block, the disc

525 and annulus alternated in anti-phase, generating a spatially segmented percept with a

526 visible boundary between the disc and annulus regions defined by relative disparity. In

527 the B block, the disc and annulus alternated in phase, leading to a uniform motion percept

528 with no border. Thus, disparity modulated between 0 and 20 arcmin at all locations in

529 both A and B blocks, with only the relative disparity over the disk and annulus regions

530 differing between A and B blocks. Participants wore anaglyph glasses throughout the

531 experiment, but the contrast and motion condition were identical in both eyes and thus

532 effectively shown at 0 disparity.

533 In Experiment 2 we replicated the relative disparity condition from Experiment 1,

534 but introduced a rapid serial visual presentation (RSVP) task at fixation that served to

535 direct attention away from the stimulus. Subjects attended to a letter F, randomly oriented

536 and superimposed on the center of the display. At random times during a block the F

537 briefly turned into either a target letter, either L or T, followed by a new F that served as

538 a mask. On each change, subjects had to indicate with a button-press whether the target

539 letter was an L or a T. The target letter duration was adapted online using a staircase

540 procedure to stabilize performance at a constant level (~80% correct) during both A and

541 B blocks.

542

543 *fMRI Experimental Procedure*

544 We used a block design in which 12 s A blocks alternated periodically with 12 s B

545 blocks, yielding a 24 s base period for the paradigm that was repeated 10 times in what

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

546   we refer to as a "scan". The design is illustrated schematically in Figure 1. Ten stimulus

547   cycles were shown per scan, with an additional half-cycle (one 12 s control block) being

548   shown in the beginning of the scan to allow the brain and the scanner to settle. The data

549   collected during this "dummy" period were removed from the fMRI time series data

550   before the data analysis. The disparity condition was not run for 2 out of 15 participants

551   in Experiment 1 because of technical issues. We acquired 4 scans per condition for each

552   participant in Experiment 1, except 3/15 participants for whom we acquired only 3 scans

553   for one or more of the conditions. In Experiment 2, we acquired 5 scans per participants,

554   except 2/10 for whom we only acquired 4 scans.

555

556   *Structural and functional MRI acquisition*

557   Functional and structural MRI data were collected on a General Electric

558   Discovery 750 (General Electric Healthcare) equipped with a 32-channel head coil (Nova

559   Medical) at the Center for Cognitive and Neurobiological Imaging at Stanford University.

560   For each participant, we acquired two whole-brain T1-weighted structural datasets (1.0 ×

561   1.0 × 1.0 mm resolution, TE=2.5 ms, TR=6.6 ms, flip angle=12, FOV=256 × 256) that

562   were used for tissue segmentation and registration with atlas-based ROIS and retinotopy

563   template (see below). For Experiment 1, a multiplexed EPI sequence [64] was used which

564   allowed for the collection of 60 horizontal slices (2.0 × 2.0 × 2.0 mm resolution, TE = 30

565   ms, TR = 2000 ms, flip angle=77, FOV=220 × 220), resulting in whole-brain coverage.

566   In Experiment 2, a non-multiplexed EPI sequence was used, which limited the coverage

567   to 32 coronal slices, positioned at an oblique angle to maximize coverage of occipital,

568   ventral and parietal cortices. The sequence used in Experiment 2 was otherwise identical

569   to the one used in Experiment 1.

570

571   *fMRI analysis*

572   After removing the dummy TRs, the fMRI data was preprocessed in AFNI [65],

573   which included the following steps: slice-time correction, motion registration (the third

574   TR of the first scan was always used as base), scaling (each voxel was scaled to a mean

575   of 100, and values were clipped at 200), and de-trending (removing components

576   corresponding to the six motion registration parameters, as well as Legendre polynomials

577   of order 0 (constant signal), 1 (linear drift) and 2).

578   The remainder of the analysis was performed in MATLAB. The time-course data

579   were first averaged across the three scans for each condition, and then across the voxels

580   within each visual region-of-interest (ROI). We then applied a Fourier transform to the

581   average time-course for each ROI, omitted DC, multiplied the spectrum by 2 to get the

582   single sided spectrum, and scaled by dividing with the number of samples in the time-

583   course. We selected the complex value at the stimulus frequency (10 cycles per scan) for

584   each participant, within each ROI, and used it for statistical analysis. For the whole-brain

585   analysis (see below), we performed the same Fourier analysis on a voxel-by-voxel basis,

586   without averaging across ROIs, which gave us a complex value at the stimulus frequency,

587   for every voxel in each participant.

588

589   *Visual regions-of-interest*

590   Topographically organized visual ROIs were derived from a probabilistic atlas[66].

591   The atlas ROIs, defined by retinotopic mapping, included 25 ROIs covering 22 visual

592   areas in ~50 individual participants. The atlas first converts the surface data from each

593   individual to surface-based standardized space, and then converts the surface data from

594   each individual to surface-based standardized space[67], finally assessing the likelihood,

595   across participants, of any particular vector on the standardized surface belonging to a

596   particular ROI[66]. The atlas was defined using a maximum probability approach, which

597   considers a given vector as part of the set of ROIs if it is more often found within the set,

598   than outside the set, across participants. If this is the case, the vector is then assigned the

599   value of the most likely ROI, and if not, it is considered to be outside the set of ROIs. The

600   maximum probability approach captures much of the overall structure of ROIs defined

601   for individual subjects and generalizes well to novel participants that did not contribute to

602   the atlas generation[66].

603   We downloaded the atlas from

604   http://scholar.princeton.edu/sites/default/files/napl/files/probatlas_v4.zip and converted

605   the ROIs from standardized surface space to native surface space for each of our

606   participants, using nearest-neighbor interpolation. We removed vertices that were more

607 than 1 edge away from the main cluster of each ROI, to ensure that all ROIs consisted

608 exclusively of contiguous vertices. This step eliminated small speckles, while having

609 minimal effect on the overall structure and extent of the ROIs. We then created a version

610 of the structural data set associated with the surface meshes that was registered to the

611 experiment data, and used that to convert the ROIs from surface space to volume space,

612 registering them to the experimental data. When multiple surface nodes were mapped to a

613 single voxel, the most common value across those nodes were assigned to the voxel.

614 Finally, the ROIs were resampled to match the resolution and extent of the experiment

615 data. We excluded four ROIs from our analysis, IPS4 and 5, TO2 and FEF, because of

616 their small size in the probabilistic atlas, and merged the dorsal and ventral segments of

617 V1, V2 and V3. This gave us a total of 18 bilateral ROIs to analyze.

618   To derive an independent estimate of the response in regions of early visual

619 cortex responding to different eccentricities in the visual field, we used a template

620 developed by Benson and colleagues[37] that accurately predicts the location and

621 retinotopic organization of early visual areas V1-V3, using only the cortical anatomy.

622 After transforming the template data to match the specific cortical topology of each

623 participant, we converted the eccentricity data in the template from surface space to

624 volume space, and registered and resampled to match the experimental data, using the

625 same approach as for the Wang ROIs. When multiple surface nodes were mapped to a

626 single voxel, the average eccentricity value across those nodes were assigned to the

627 voxel. We could now sub-divide the ROIs in early visual cortex, for each participant, by

628 selecting voxels within V1, V2 and V3 that were responsive to a given range of

629 eccentricities. We generated 24 sub-ROIs for each early visual area, centered on radii

630 ranging from 0.25 to 6.0º, separated by 0.25º, and each spanning 0.5º/vis. angle.

631

632 *Vector-based statistics*

633   We computed the average phase and amplitude at the stimulus frequency using a

634 vector-based approach, in which the real and imaginary part of the complex value was

635 averaged separately across participants, and then combined so that vector mean amplitude

636 and phase could be computed. Error bars were computed using a geometrical approach,

637 in which a two-dimensional error ellipse is computed, which describes the standard error

638 of the mean response amplitude. The upper and lower error bounds were computed as the

639 longest and shortest vectors from the origin to the error ellipse a detailed describtion of

640 this approach can be found in [68]. All statistical tests for significance were run as

641 Hotelling's t^2 tests of the null hypothesis that the two-dimensional data set containing the

642 real and imaginary parts of the complex value at the stimulus frequency was equal to

643 [0,0][69]. Note that this vector-based approach means that both amplitude and phase, and

644 their consistency across participants, contributes to our reported estimates of mean

645 amplitude, error and statistical significance.

646 We computed the sign of the responses by doing a linear fit with zero intercept of

647 the real and imaginary values associated with the contrast condition, averaged across

648 participants, within each of the eccentricity-based sub-divisions of V1. The amplitude of

649 the response to the contrast condition was high at most retinotopic locations, but response

650 phase varied with eccentricity. Values that were to the left and below a line orthogonal to

651 the fit line were given negative phase signs (weaker responses to the A block than B

652 block), while values above and to the right were given positive phase signs (stronger

653 responses to the A block than B block). The contrast-based fit was used for all conditions.

654 We multiplied mean amplitude with phase sign in the ROI plots, to illustrate the phase of

655 the ROIs response.

656

657 *Whole-brain analysis*

658 To provide an overview of the effect of our conditions across the whole brain, and

659 account for any potential effects outside our set of visual ROIs, we mapped the complex

660 values at each voxel onto a standardized cortical surface, for each participant. This

661 surface-based alignment offers several advantages over volume-based approaches to

662 group analysis[70], most importantly by considering the structure of cortical sulci and gyri,

663 as opposed to Talairach registration and other types cross-subject normalization in

664 volume-space, which is likely to blur activations across neighboring banks of a sulcus[71].

665 After surface-based alignment, we used the real and imaginary parts of the complex value

666 at the stimulus frequency, at each surface node across participants, to compute mean

667 amplitude and phase using the same vector-based approach applied to the ROI data. We

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

668 also performed the Hotelling's t^2 test for significance as described above, for each node

669 on the surface, and used that for thresholding the surface data.

670

671 **Acknowledgements**

672 This research was supported by NIH grant EY018775.

673

674 **Author Contributions**

675 PJK, BRC and AMN designed the study. PJK and BRC collected the data. PJK analyzed

676 the data. PJK, BRC and AMN wrote the paper.

677

678 **Additional Information**

679 The authors declare no competing financial interests.
680
681
682 **Figure Legends**

683 **Figure 1. The experimental design used in Experiment 1.** The on-off responses evoked

684 by the conditions that are captured by our Fourier analysis, are illustrated with sine

685 waves. (A) The contrast condition – the orientation of both the disc and annulus

686 alternated between vertical and horizontal at 2 Hz. Disk and annulus were presented in

687 temporal succession. (B) The motion condition – dots alternated between left- and right-

688 wards motion at 2Hz. When the disc and annulus regions alternated in anti-phase, a figure

689 percept was evoked, when alternations were in-phase, there was no such percept. (C) The

690 disparity condition – the dots alternated between the fixation plane (0 arcmin) and a

691 position behind fixation (uncrossed disparity, 20 arcmin) at 2 Hz. Anti-phase movement

692 of the disc and annulus regions lead to a figure percept, while in-phase movement

693 generated no such percept. The white dotted line in B and C indicates the extent of the

694 disc and annulus used in A. Note that the relative disparity condition was identical to that

695 used in Experiment 2, except for the introduction of an RSVP task at fixation, which

696 directed attention away from the stimulus. (D) Schematic outlining the size of the disc

697 and annulus region, respectively, which was shared across all conditions.

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

698

699 **Figure 2: Eccentricity analysis of V1-V3 in Experiment 1.** Signed vector mean

700 amplitudes for the contrast (A), motion (B) and disparity (C) conditions within 24

701 eccentricity-defined sub-ROIs, centered on eccentricities spaced 0.25º apart, each having

702 a width of 0.5º. The shaded areas on the plots indicate condition ✕ sub-ROI

703 combinations that were significant at $\alpha = 0.05$. The color of the text for the ROI names

704 matches the ROI colors on Figures 4 and 6.

705

706 **Figure 3: Eccentricity analysis of V1-V3 in Experiment 2.** Signed vector mean

707 amplitudes within 24 eccentricity-defined sub-ROIs, centered on eccentricities spaced

708 0.25º apart, each having a width of 0.5º. The shaded areas on the plots indicate condition

709 ✕ sub-ROI combinations that were significant at $\alpha = 0.05$. The color of the text for the

710 ROI names matches the ROI colors on Figures 4 and 6. The relative disparity condition

711 shown here was identical to the one used in Experiment 1 (see Figure 2C), except

712 attention was directed away from the stimulus.

713

714 **Figure 4: ROI results from Experiment 1.** Signed vector mean amplitude within 15

715 topographically organized regions of interest, excluding early visual areas V1-V3, which

716 are shown in Figure 2. The motion condition is shown in orange, and the disparity

717 condition in green. Shaded areas behind the bars indicate condition × ROI combinations

718 that were significant at $\alpha = 0.05$. The ROIs are shown for the both hemispheres of an

719 example participant's inflated cortical surface reconstruction, below the graph.

720

721 **Figure 5: ROI results from Experiment 2.** Signed vector mean amplitude within 12

722 topographically organized regions of interest, excluding early visual areas V1-V3, which

723 are shown in Figure 3, and three ROIs that were not covered by the fMRI acquisition

724 protocol used in Experiment 2 (VO2, PHC1 and PHC2). Shared areas behind the bars

725 indicate condition × ROI combinations that were significant at $\alpha = 0.05$. The ROIs are

726 shown for the right hemisphere of an example participant's inflated cortical surface

727 reconstruction, on the right side of the graph.

728

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

729 **Figure 6: Whole-brain results from Experiment 1.** Vector-averaged phase maps,

730 thresholded at $\alpha = 0.05$. Maps were produced through surface-based alignment procedure

731 in which each subjects' cortical mesh was converted to a standardized mesh, which

732 allowed for cross-subject comparisons of values at each mesh node. The vector average

733 phase across subjects, as well as a corresponding *p*-value based on both amplitude and

734 phase, could then be computed for each mesh node. Blue shades indicate responses to the

735 A block, while orange shades indicate responses to the B block. The ROIs are outlined on

736 the surface, and labeled version can be inspected in Figure 4 and 5.

737

738 **References**
739

740 1   Ban, H., Preston, T. J., Meeson, A. & Welchman, A. E. The integration of motion
741   and disparity cues to depth in dorsal visual cortex. *Nature neuroscience* **15**, 636-
742   643, doi:http://www.nature.com/neuro/journal/v15/n4/abs/nn.3046.html -
743   supplementary-information (2012).
744 2   Bradshaw, M. F. & Rogers, B. J. The interaction of binocular disparity and
745   motion parallax in the computation of depth. *Vision research* **36**, 3457-3468
746   (1996).
747 3   Nawrot, M. & Blake, R. Neural integration of information specifying structure
748   from stereopsis and motion. *Science* **244**, 716-718, doi:10.1126/science.2717948
749   (1989).
750 4   Lamme, V. A. F., van Dijk, B. W. & Spekreijse, H. Contour from motion
751   processing occurs in primary visual cortex. *Nature* **363**, 541-543 (1993).
752 5   Cao, A. N. & Schiller, P. H. Neural responses to relative speed in the primary
753   visual cortex of rhesus monkey. *Visual neuroscience* **20**, 77-84,
754   doi:10.1017/S0952523803201085 (2003).
755 6   Lui, L. L., Bourne, J. A. & Rosa, M. G. P. Single-unit responses to kinetic stimuli
756   in New World monkey area V2: Physiological characteristics of cue-invariant
757   neurones. *Experimental Brain Research* **162**, 100-108, doi:10.1007/s00221-004-
758   2113-9 (2005).
759 7   Shen, Z.-M., Xu, W.-F. & Li, C.-Y. Cue-invariant detection of centre–surround
760   discontinuity by V1 neurons in awake macaque monkey. *The Journal of
761   Physiology* **583**, 581-592, doi:10.1113/jphysiol.2007.130294 (2007).
762 8   Yin, J. *et al.* Breaking cover: neural responses to slow and fast camouflage-
763   breaking motion. *Proceedings of the Royal Society B: Biological Sciences* **282**,
764   doi:10.1098/rspb.2015.1182 (2015).
765 9   Chen, M. *et al.* An Orientation Map for Motion Boundaries in Macaque V2.
766   *Cerebral cortex* **26**, 279-287, doi:10.1093/cercor/bhu235 (2016).
767 10   Allman, J., Miezin, F. & McGuinness, E. Direction- and Velocity-Specific
768   Responses from beyond the Classical Receptive Field in the Middle Temporal
769   Visual Area (MT). *Perception* **14**, 105-126, doi:doi:10.1068/p140105 (1985).

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW

770   11   Marcar, V. L., Xiao, D. K., Raiguel, S. E., Maes, H. & Orban, G. A. Processing of
771   kinetically defined boundaries in the cortical motion area MT of the macaque
772   monkey. *J Neurophysiol* **74**, 1258-1270 (1995).
773   12   Raiguel, S., Van Hulle, M. M., Xiao, D. K., Marcar, V. L. & Orban, G. A. Shape
774   and spatial distribution of receptive fields and antagonistic motion surrounds in
775   the middle temporal area (V5) of the macaque. *The European journal of*
776   *neuroscience* **7**, 2064-2082 (1995).
777   13   Xiao, D. K., Raiguel, S., Marcar, V. & Orban, G. A. The spatial distribution of the
778   antagonistic surround of MT/V5 neurons. *Cereb Cortex* **7**, 662-677 (1997).
779   14   Sary, G., Vogels, R., Kovacs, G. & Orban, G. A. Responses of monkey inferior
780   temporal neurons to luminance-, motion-, and texture-defined gratings. *Journal of*
781   *neurophysiology* **73**, 1341-1354 (1995).
782   15   Marcar, V. L., Raiguel, S. E., Xiao, D. & Orban, G. A. Processing of Kinetically
783   Defined Boundaries in Areas V1 and V2 of the Macaque Monkey. *Journal of*
784   *neurophysiology* **84**, 2786-2798 (2000).
785   16   Dupont, P. *et al.* The kinetic occipital region in human visual cortex. *Cereb*
786   *Cortex* **7**, 283-292 (1997).
787   17   Van Oostende, S., Sunaert, S., Van Hecke, P., Marchal, G. & Orban, G. A. The
788   kinetic occipital (KO) region in man: an fMRI study. *Cerebral cortex* **7**, 690-701
789   (1997).
790   18   Reppas, J. B., Niyogi, S., Dale, A. M., Sereno, M. I. & Tootell, R. B. H.
791   Representation of motion boundaries in retinotopic human visual cortical areas.
792   *Nature* **388**, 175-179 (1997).
793   19   Shulman, G. L., Schwarz, J., Miezin, F. M. & Petersen, S. E. Effect of Motion
794   Contrast on Human Cortical Responses to Moving Stimuli. *Journal of*
795   *neurophysiology* **79**, 2794-2803 (1998).
796   20   Skiera, G., Petersen, D., Skalej, M. & Fahle, M. Correlates of figure-ground
797   segregation in fMRI. *Vision research* **40**, 2047-2056,
798   doi:https://doi.org/10.1016/S0042-6989(00)00038-9 (2000).
799   21   Larsson, J., Heeger, D. J. & Landy, M. S. Orientation Selectivity of Motion-
800   Boundary Responses in Human Visual Cortex. *Journal of neurophysiology* **104**,
801   2940-2950, doi:10.1152/jn.00400.2010 (2010).
802   22   Thomas, O. M., Cumming, B. G. & Parker, A. J. A specialization for relative
803   disparity in V2. *Nature neuroscience* **5**, 472-478, doi:10.1038/nn837 (2002).
804   23   Bredfeldt, C. E. & Cumming, B. G. A simple account of cyclopean edge
805   responses in macaque v2. *J Neurosci* **26**, 7581-7596,
806   doi:10.1523/JNEUROSCI.5308-05.2006 (2006).
807   24   Qiu, F. T. & von der Heydt, R. Figure and ground in the visual cortex: v2
808   combines stereoscopic cues with gestalt rules. *Neuron* **47**, 155-166,
809   doi:10.1016/j.neuron.2005.05.028 (2005).
810   25   Anzai, A., Chowdhury, S. A. & DeAngelis, G. C. Coding of Stereoscopic Depth
811   Information in Visual Areas V3 and V3A. *The Journal of Neuroscience* **31**,
812   10270-10282, doi:10.1523/jneurosci.5956-10.2011 (2011).
813   26   Umeda, K., Tanabe, S. & Fujita, I. Representation of Stereoscopic Depth Based
814   on Relative Disparity in Macaque Area V4. *Journal of neurophysiology* **98**, 241-
815   252, doi:10.1152/jn.01336.2006 (2007).

816 27 Shiozaki, H. M., Tanabe, S., Doi, T. & Fujita, I. Neural Activity in Cortical Area
817   V4 Underlies Fine Disparity Discrimination. *The Journal of Neuroscience* **32**,
818   3830-3841, doi:10.1523/jneurosci.5083-11.2012 (2012).
819 28 Janssen, P., Vogels, R., Liu, Y. & Orban, G. A. Macaque Inferior Temporal
820   Neurons Are Selective for Three-Dimensional Boundaries and Surfaces. *The*
821   *Journal of Neuroscience* **21**, 9419-9429 (2001).
822 29 Krug, K. & Parker, A. J. Neurons in Dorsal Visual Area V5/MT Signal Relative
823   Disparity. *The Journal of Neuroscience* **31**, 17892-17904,
824   doi:10.1523/jneurosci.2658-11.2011 (2011).
825 30 Tsao, D. Y. *et al.* Stereopsis Activates V3A and Caudal Intraparietal Areas in
826   Macaques and Humans. *Neuron* **39**, 555-568, doi:https://doi.org/10.1016/S0896-
827   6273(03)00459-8 (2003).
828 31 Backus, B. T., Fleet, D. J., Parker, A. J. & Heeger, D. J. Human Cortical Activity
829   Correlates With Stereoscopic Depth Perception. *Journal of neurophysiology* **86**,
830   2054-2068 (2001).
831 32 Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K. & Tootell, R. B. H. The
832   Representation of Illusory and Real Contours in Human Cortical Visual Areas
833   Revealed by Functional Magnetic Resonance Imaging. *The Journal of*
834   *Neuroscience* **19**, 8560-8572 (1999).
835 33 Neri, P., Bridge, H. & Heeger, D. J. Stereoscopic Processing of Absolute and
836   Relative Disparity in Human Visual Cortex. *Journal of neurophysiology* **92**, 1880-
837   1891, doi:10.1152/jn.01042.2003 (2004).
838 34 Vinberg, J. & Grill-Spector, K. Representation of Shapes, Edges, and Surfaces
839   Across Multiple Cues in the Human Visual Cortex. *Journal of neurophysiology*
840   **99**, 1380-1393, doi:10.1152/jn.01223.2007 (2008).
841 35 Cottereau, B. R., McKee, S. P., Ales, J. M. & Norcia, A. M. Disparity-Tuned
842   Population Responses from Human Visual Cortex. *The Journal of Neuroscience*
843   **31**, 954-965, doi:10.1523/jneurosci.3795-10.2011 (2011).
844 36 Cottereau, B. R., McKee, S. P., Ales, J. M. & Norcia, A. M. Disparity-Specific
845   Spatial Interactions: Evidence from EEG Source Imaging. *The Journal of*
846   *Neuroscience* **32**, 826-840, doi:10.1523/jneurosci.2709-11.2012 (2012).
847 37 Benson, N. C., Butt, O. H., Brainard, D. H. & Aguirre, G. K. Correction of
848   Distortion in Flattened Representations of the Cortical Surface Allows Prediction
849   of V1-V3 Functional Organization from Anatomy. *PLoS computational biology*
850   **10**, e1003538, doi:10.1371/journal.pcbi.1003538 (2014).
851 38 Gandhi, S. P., Heeger, D. J. & Boynton, G. M. Spatial attention affects brain
852   activity in human primary visual cortex. *Proceedings of the National Academy of*
853   *Sciences* **96**, 3314-3319 (1999).
854 39 Martinez, A. *et al.* Involvement of striate and extrastriate visual cortical areas in
855   spatial attention. *Nature neuroscience* **2**, 364-369 (1999).
856 40 Somers, D. C., Dale, A. M., Seiffert, A. E. & Tootell, R. B. Functional MRI
857   reveals spatially specific attentional modulation in human primary visual cortex.
858   *Proceedings of the National Academy of Sciences* **96**, 1663-1668 (1999).
859 41 Parker, A. J. Binocular depth perception and the cerebral cortex. *Nature reviews*
860   **8**, 379-391 (2007).

861 42   Neri, P. A stereoscopic look at visual cortex. *J Neurophysiol* **93**, 1823-1826
862   (2005).
863 43   Verhoef, B. E., Vogels, R. & Janssen, P. Binocular depth processing in the ventral
864   visual pathway. *Philosophical transactions of the Royal Society of London* **371**,
865   doi:10.1098/rstb.2015.0259 (2016).
866 44   Tanabe, S., Umeda, K. & Fujita, I. Rejection of false matches for binocular
867   correspondence in macaque visual cortical area V4. *J Neurosci* **24**, 8170-8180
868   (2004).
869 45   Mysore, S. G., Vogels, R., Raiguel, S. E. & Orban, G. A. Processing of Kinetic
870   Boundaries in Macaque V4. *Journal of neurophysiology* **95**, 1864-1880,
871   doi:10.1152/jn.00627.2005 (2006).
872 46   Uka, T., Tanabe, S., Watanabe, M. & Fujita, I. Neural correlates of fine depth
873   discrimination in monkey inferior temporal cortex. *J Neurosci* **25**, 10796-10802
874   (2005).
875 47   Verhoef, B. E., Vogels, R. & Janssen, P. Inferotemporal cortex subserves three-
876   dimensional structure categorization. *Neuron* **73**, 171-182,
877   doi:10.1016/j.neuron.2011.10.031 (2012).
878 48   Unno, S., Handa, T., Nagasaka, Y., Inoue, M. & Mikami, A. Modulation of
879   neuronal activity with cue-invariant shape discrimination in the primate superior
880   temporal sulcus. *Neuroscience* **268**, 221-235,
881   doi:10.1016/j.neuroscience.2014.03.024 (2014).
882 49   Haak, K. V. & Beckmann, C. F. Objective analysis of the topological organization
883   of the human cortical visual connectome suggests three visual pathways. *Cortex;*
884   *a journal devoted to the study of the nervous system and behavior*,
885   doi:10.1016/j.cortex.2017.03.020 (2017).
886 50   Weiner, K. S. & Grill-Spector, K. Neural representations of faces and limbs
887   neighbor in human high-level visual cortex: evidence for a new organization
888   principle. *Psychol Res* **77**, 74-97, doi:10.1007/s00426-011-0392-x (2013).
889 51   Saad, Z. S., DeYoe, E. A. & Ropella, K. M. Estimation of FMRI response delays.
890   *NeuroImage* **18**, 494-504 (2003).
891 52   Shmuel, A. *et al.* Sustained negative BOLD, blood flow and oxygen consumption
892   response and its coupling to the positive response in the human brain. *Neuron* **36**,
893   1195-1210 (2002).
894 53   Shmuel, A., Augath, M., Oeltermann, A. & Logothetis, N. K. Negative functional
895   MRI response correlates with decreases in neuronal activity in monkey visual area
896   V1. *Nature neuroscience* **9**, 569-577, doi:10.1038/nn1675 (2006).
897 54   Dougherty, R. F. *et al.* Visual field representations and locations of visual areas
898   V1/2/3 in human visual cortex. *Journal of vision* **3**, 586-598, doi:10:1167/3.10.1
899   (2003).
900 55   Bridge, H. & Parker, A. J. Topographical representation of binocular depth in the
901   human visual cortex using fMRI. *Journal of vision* **7**, 15-15, doi:10.1167/7.14.15
902   (2007).
903 56   Likova, L. T. & Tyler, C. W. Occipital network for figure/ground organization.
904   *Experimental Brain Research* **189**, 257, doi:10.1007/s00221-008-1417-6 (2008).
905 57   Cacciamani, L., Scalf, P. E. & Peterson, M. A. Neural evidence for competition-
906   mediated suppression in the perception of a single object. *Cortex; a journal*

907   *devoted to the study of the nervous system and behavior* **72**, 124-139,
908   doi:10.1016/j.cortex.2015.05.018 (2015).
909   58   Amano, K., Wandell, B. A. & Dumoulin, S. O. Visual Field Maps, Population
910   Receptive Field Sizes, and Visual Field Coverage in the Human MT+ Complex.
911   *Journal of neurophysiology* **102**, 2704-2718, doi:10.1152/jn.00102.2009 (2009).
912   59   Harvey, B. M. & Dumoulin, S. O. The relationship between cortical
913   magnification factor and population receptive field size in human visual cortex:
914   constancies in cortical architecture. *The Journal of neuroscience : the official*
915   *journal of the Society for Neuroscience* **31**, 13604-13612,
916   doi:10.1523/JNEUROSCI.2572-11.2011 (2011).
917   60   Dumoulin, S. O. & Wandell, B. A. Population receptive field estimates in human
918   visual cortex. *NeuroImage* **39**, 647-660, doi:10.1016/j.neuroimage.2007.09.034
919   (2008).
920   61   Salvagio, E., Cacciamani, L. & Peterson, M. A. Competition-strength-dependent
921   ground suppression in figure-ground perception. *Attention, perception &*
922   *psychophysics* **74**, 964-978, doi:10.3758/s13414-012-0280-5 (2012).
923   62   Poort, J., Self, M. W., van Vugt, B., Malkki, H. & Roelfsema, P. R. Texture
924   Segregation Causes Early Figure Enhancement and Later Ground Suppression in
925   Areas V1 and V4 of Visual Cortex. *Cerebral cortex* **26**, 3964-3976,
926   doi:10.1093/cercor/bhw235 (2016).
927   63   Qiu, F. T., Sugihara, T. & von der Heydt, R. Figure-ground mechanisms provide
928   structure for selective attention. *Nature neuroscience* **10**, 1492-1499 (2007).
929   64   Feinberg, D. A. *et al.* Multiplexed echo planar imaging for sub-second whole
930   brain FMRI and fast diffusion imaging. *PloS one* **5**, e15710,
931   doi:10.1371/journal.pone.0015710 (2010).
932   65   Cox, R. W. AFNI: software for analysis and visualization of functional magnetic
933   resonance neuroimages. *Computers and Biomedical research* **29**, 162-173 (1996).
934   66   Wang, H. X., Merriam, E. P., Freeman, J. & Heeger, D. J. Motion Direction
935   Biases and Decoding in Human Visual Cortex. *The Journal of Neuroscience* **34**,
936   12601-12615, doi:10.1523/jneurosci.1034-14.2014 (2014).
937   67   Argall, B. D., Saad, Z. S. & Beauchamp, M. S. Simplified intersubject averaging
938   on the cortical surface using SUMA. *Human brain mapping* **27**, 14-27,
939   doi:10.1002/hbm.20158 (2006).
940   68   Pei, F., Baldassi, S., Tsai, J. J., Gerhard, H. E. & Norcia, A. M. Development of
941   contrast normalization mechanisms during childhood and adolescence. *Vision*
942   *research* **133**, 12-20, doi:https://doi.org/10.1016/j.visres.2016.03.010 (2017).
943   69   Anderson, T. W. *An introduction to multivariate statistical analysis*. (Wiley,
944   1984).
945   70   Hagler, D. J., Saygin, A. P. & Sereno, M. I. Smoothing and cluster thresholding
946   for cortical surface-based group analysis of fMRI data. *NeuroImage* **33**, 1093-
947   1103, doi:https://doi.org/10.1016/j.neuroimage.2006.07.036 (2006).
948   71   Fischl, B., Sereno, M. I. & Dale, A. M. Cortical surface-based analysis: II:
949   Inflation, flattening, and a surface-based coordinate system. *NeuroImage* **9**, 195-
950   207 (1999).
951

THIS MANUSCRIPT HAS NOT BEEN SUBJECT TO PEER REVIEW