1 **Body shape as a visual feature: evidence from spatially-global attentional modulation in**

2 **human visual cortex**

3

4 Abbreviated title: Body shape as a visual feature

5

6 Sushrut Thorat[1], Marius V. Peelen[1]

7

8 [1]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The

9 Netherlands.

10

11 Corresponding author:

12 Marius V. Peelen

13 Donders Institute for Brain, Cognition and Behaviour

14 Radboud University

15 Thomas van Aquinostraat 4, 6525 GD Nijmegen, The Netherlands

16 Email: marius.peelen@donders.ru.nl

17

18

19

20 Conflict of interest statement: The authors declare no competing financial interests.

Abstract

Feature-based attention modulates visual processing beyond the focus of spatial attention.

Previous work has reported such spatially-global effects for low-level features such as color and

orientation, as well as for faces. Here, using fMRI, we provide evidence for spatially-global

attentional modulation for human bodies. Participants were cued to search for one of six object

categories in two vertically-aligned images. Two additional, horizontally-aligned, images were

simultaneously presented but were never task-relevant across three experimental sessions.

Analyses time-locked to the objects presented in these task-irrelevant images revealed that

responses evoked by body silhouettes were modulated by the participants' top-down attentional

set, becoming more body-selective when participants searched for bodies in the task-relevant

images. These effects were observed both in univariate analyses of the body-selective cortex

and in multivariate analyses of the object-selective visual cortex. Additional analyses showed

that this modulation reflected response gain rather than a bias induced by the cues, and that it

reflected enhancement of body responses rather than suppression of non-body responses.

These findings provide evidence for a spatially-global attention mechanism for body shapes,

supporting the rapid and parallel detection of conspecifics in our environment.

Introduction

The capacity limits of the human visual system require selecting visual input for further processing and conscious access (Carrasco, 2011; Chun et al., 2011). One way to do this is to select specific locations of the visual field through spatial attention and eye movements. However, when searching for task-relevant objects in our environment, the location of these objects is typically not yet known. In this case, selection may operate at the level of visual features, using a selection mechanism termed feature-based attention (Maunsell and Treue, 2006). To be an effective selection mechanism, feature-based attention would need to operate in parallel across the whole or part of the visual field, in order to then guide spatial attention to the location of the target object (Wolfe, 1994). While this could be a plausible mechanism of attentional selection, it raises a core question: what are the features of feature-based attention?

At a neural level, it has been proposed that feature-based attention may be restricted to features to which sensory neurons are systematically tuned (Maunsell and Treue, 2006). Accordingly, the neural mechanisms of feature-based attention have been studied extensively with experiments involving low-level features for which such tuning has been established, such as the orientations of Gabor patches (Kamitani and Tong, 2005; Liu et al., 2007; Jehee et al., 2011) and the movement direction of random dot patterns (Treue and Trujillo, 1999; Saenz et al., 2002; Serences and Boynton, 2007). These experiments assessed how making one feature task-relevant influenced the responses of neurons that were selective or non-selective to that feature. A common finding was that attending to a low-level feature increased the responses of neurons selective to that feature and decreased the responses of neurons non-selective to that feature (Maunsell and Treue, 2006). Crucially, such modulations were shown to occur for stimuli presented in spatially-unattended and task-irrelevant locations (Treue and Trujillo, 1999; Saenz et al., 2002; Serences and Boynton, 2007; Zhang and Luck, 2009), providing evidence for a

64 spatially-global mechanism of feature-based attention that can be distinguished from the effects

65 of spatial attention.

66   In the present study, we tested whether global attentional modulation can similarly be

67 observed for the shape of the human body, a category of high social and biological significance

68 that is selectively represented in high-level visual cortex (Downing et al., 2001; Peelen &

69 Downing, 2005). Behavioral studies have shown that bodies, like faces, gain preferential access

70 to awareness (Stein et al., 2012) and automatically attract attention (Downing et al., 2004; Ro et

71 al., 2007). There is also behavioral evidence for spatially-global attention effects for bodies: in a

72 series of studies, spatial attention was captured by body silhouettes when participants searched

73 for people in scenes presented in different parts of the visual field (Reeder and Peelen, 2013;

74 Reeder et al., 2015). Finally, an fMRI study reported spatially-global modulation of multivoxel

75 activity patterns distinguishing natural scenes with people from natural scenes with cars (Peelen

76 et al., 2009). However, in that study, the relative contributions of scene context and of body,

77 face, and car features could not be distinguished, such that it remains unknown whether

78 feature-based attention effects exist for human bodies.

79   Here, we used fMRI to test for spatially-global attentional modulation of body processing

80 in visual cortex. Participants detected the presence of bodies or one of five other categories

81 (beds, bottles, cars, chairs, lamps) in task-relevant vertically-aligned images, thereby

82 manipulating the top-down attentional set. To test for spatially-global attentional modulation, all

83 analyses focused on responses evoked by objects that were concurrently presented at locations

84 that were never relevant for the object detection task across three experimental sessions (Fig.

85 1A). The inclusion of five non-body categories reduced the possibility that participants could use

86 a low-level feature to detect the presence of bodies, for example by looking for vertical (bodies)

87 vs horizontal (e.g., cars) stimuli: lamps and bottles shared the vertical orientation with bodies

88 (Fig 1D). To further reduce this possibility, each category was represented by a large and

89 diverse set of exemplars cropped out of scene photographs. Finally, the use of silhouettes

4

90   avoided possible low-level differences between categories in texture and color, and ensured

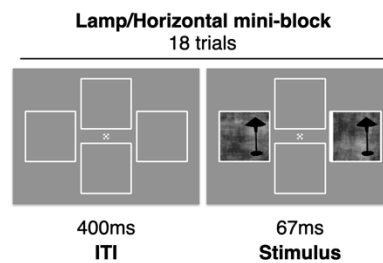91   that attention was guided by body shape rather than facial features (Störmer et al., 2019).
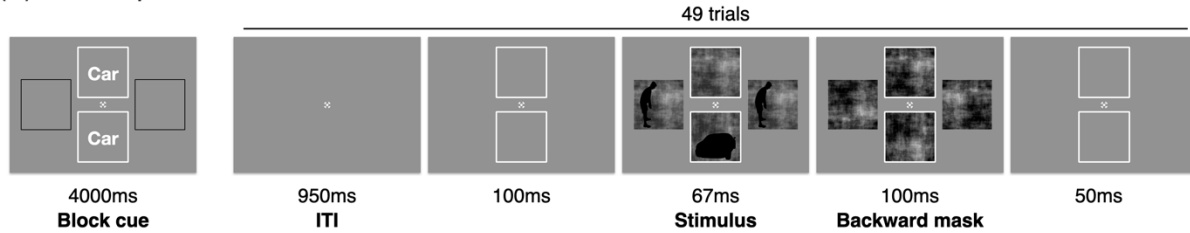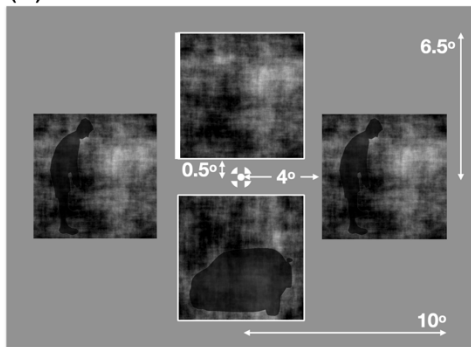
Figure 1: Experimental design. (A) The main experiment was designed to reveal the modulatory influence of feature-based attention on object responses evoked by stimuli presented at task-irrelevant locations (horizontal boxes). In each block (49 trials), participants had to search for the cued object category (e.g., car) in the vertical boxes, while objects were simultaneously presented in the horizontal boxes. (B) The spatial layout of the search display. (C) The baseline experiment was included to localize body-selective regions of interest (for the univariate analyses) and to obtain prototypical object category response patterns (for the multivariate analyses). Responses evoked by task-irrelevant objects in the main experiment were compared to these responses. Participants had to indicate if one of the edges of the two boxes thickened. The object category and location (horizontal or vertical boxes) varied across the mini-blocks. Unlike in the main experiment, the stimuli were not backward masked in order to increase visibility. (D) Exemplars of the six object categories: chairs, lamps, beds, cars, human bodies, bottles. Fifty exemplars were used for each category.

Materials and Methods

Participants

Twenty-three healthy adult volunteers with normal or corrected-to-normal vision gave written informed consent and participated in the experiment. All participants took part in three experimental sessions, on different days. One participant was excluded because of low performance on the visual search task (the difference between the proportion of false alarms and hits was lower than two standard deviations from the average difference). Twenty-two participants (mean age: 25.36 years; age range: 20-32 years; 11 female) were included in the reported analyses. The study was approved by the local ethics committee (CMO Arnhem-Nijmegen).

Experimental Paradigm

In the main experiment, on each trial, the display contained two boxes in the horizontal and vertical locations (Fig. 1). The vertical boxes had a white bounding frame, signifying their relevance. Each of the four boxes contained a random image containing the average power spectrum of the objects from the six categories with random phases. Objects were mixed with these random images. On each trial, an exemplar from one of the six categories could be presented in one of the two vertical boxes (1/7 probability each) or no object would be presented (1/7 probability). Simultaneously, an exemplar from one of the six categories could be presented in both the horizontal boxes (1/7 probability each) or no object would be presented (1/7 probability). Each block consisted of 49 trials to fill the co-occurrence matrix of the horizontal and vertical object conditions, such that the conditions presented in the horizontal and vertical boxes were orthogonal to each other.

In each block of the main experiment, participants would either search for one of the six categories in the vertical boxes or would detect a thickening of the frames of the bounding

6

130   boxes in the vertical location (for trial layout, see Fig. 1). Participants pressed the response

131   button when the cued object category was shown in one of the vertical locations, which

132   occurred on 7/49 trials. Participants had to respond within 1.2s. The brief presentation duration

133   (67 ms) required participants to maintain fixation to be able to detect the target in one of the two

134   vertical locations. Participants were instructed that they could ignore the objects presented at

135   the horizontal locations. In the thickening condition, participants had to indicate, by pressing the

136   response button, when one of the sides of the two bounding boxes became thicker than the

137   others (thickening occurred on 7/49 trials in all blocks). Data from these thickening task blocks

138   in the main experiment were not further analyzed because the within-block comparisons

139   provided a more stringent test of our hypotheses, controlling for block-based effects (e.g.,

140   related to the processing of the category cue itself). The simultaneously presented objects in the

141   horizontal boxes were always task-irrelevant. Each run contained four blocks, all containing a

142   different search condition, such that across the seven search runs in each fMRI session each

143   search block occurred four times. Feedback about search performance was provided at the end

144   of each block.

145   In addition to the main experiment, participants completed a "baseline" experiment. This

146   experiment was included to localize body-selective regions of interest (for the univariate

147   analyses) and to obtain prototypical object category response patterns (for the multivariate

148   analyses). In the baseline experiment, in different blocks, exemplars of one of the six categories

149   or scrambled exemplars of one of the six categories were presented in both the boxes in either

150   the horizontal or vertical locations (the other location left empty). These objects were mixed with

151   a random image containing the average power spectrum of the objects from the six categories

152   with random phases. The seven object conditions (six object categories and a scrambled

153   objects condition containing a mix of scrambled objects from the six categories) and two

154   presentation locations were blocked into mini-blocks containing 18 trials each. In each mini-

155   block, participants had to search for thickening of the frames of the boxes where objects were

7

156   being presented (1/7 probability of presence; each pair of thickening events had at least two

157   non-thickening trials between them). Each block contained seven mini-blocks, with distinct

158   object-location pairing, such that across the four blocks in each baseline experiment run, each

159   type of block occurred twice. At the end of each block performance feedback was provided.

160   Each participant attended three experimental sessions. The first behavioral session

161   required each participant to get exposed to the entire set of objects followed by the completion

162   of one run of the baseline experiment and two runs of the main experiment. The second and the

163   third sessions involved fMRI. In each of those sessions, the participant first browsed through the

164   entire set of objects at their own pace, and then performed one run of the main experiment

165   during the anatomical scan. This was followed by the functional recordings as the participants

166   performed one run of the baseline experiment followed by four runs of the main experiment

167   followed by one run of the baseline experiment followed by three runs of the main experiment.

168   *Stimuli*

169   The stimulus presentation dimensions are shown in Fig. 1B. We acquired 50 exemplar

170   silhouettes in real-world poses for each of the six categories of interest (beds, bottles, cars,

171   chairs, lamps, and human bodies; shown in Fig. 1D). We obtained scenes containing the

172   relevant objects from the SUN2012 database (Xiao et al., 2010) and Google images which were

173   "Labelled for non-commercial reuse with modifications", cropped out the objects, scaled them

174   such than on one of the axes of the objects extended throughout the image, and converted

175   them to silhouettes.

176   On each trial, the chosen exemplars were shown in the boxes, embedded in noise as

177   mentioned above. The location of the objects within the boxes was jittered to increase

178   variability. Objects that extended throughout the image horizontally were presented in one of

179   three places within the box: touching the upper side, centered, or touching the lower side of the

180   box. Similarly, objects that extended throughout the image vertically could be placed touching

181   the left side, centered, or the right side of the box. The horizontally-placed boxes in the display

182   contained the same stimulus (Fig. 1C).

183   *fMRI data acquisition and preprocessing*

184   Functional (echo-planar imaging (EPI) sequence; 66 transverse slices per volume; resolution:

185   2x2x2mm; repetition time (TR): 1s; time to echo (TE): 35.2ms; flip angle: 60°; 6x multi-band

186   acceleration factor) and anatomical (MPRAGE sequence; 192 sagittal slices; TR: 2.3s; TE:

187   3.03ms; flip angle: 8°; 1x1x1mm resolution; FOV: 256mm) images were acquired with a 3T

188   MAGNETOM Skyra MR scanner (Siemens AG, Healthcare Sector, Erlangen, Germany) using a

189   32-channel head coil.

190   The functional data were analyzed using MATLAB (2017a) and SPM12. During

191   preprocessing, within each session, the functional volumes were realigned, co-registered to the

192   structural image, re-sampled to a 2×2×2mm grid, and spatially normalized to the Montreal

193   Neurological Institute 305 template included in SPM12. Data were high-pass filtered with a cut-

194   off of 128s. Temporal autocorrelations were accounted for using the AR(1) method in SPM. A

195   gaussian filter (FWHM 3 mm) was applied to smooth the images.

196   *Statistical analysis*

197   For each participant, general linear models (GLMs) were created to model the conditions in the

198   experiment. In the main experiment, the GLM included regressors for the 49 conditions of

199   interest: 7 attention blocks x 7 stimulus conditions presented in the task-irrelevant (horizontal)

200   locations. As this was an event-related design, the onsets of the stimuli were modelled as

201   impulse functions (delta functions) and the time series was convolved with the canonical HRF.

202   In the baseline experiment, the GLM included regressors for the 14 conditions of interest: 7

203   stimulus conditions x 2 presentation locations. As this was a block-design, the mini-blocks

204   corresponding to each stimulus condition were modelled as boxcars and the time series was

9

205   convolved with the canonical HRF. Separate GLMs were executed for each run of the main and

206   baseline experiments. The acquired regression weights were averaged across repetitions of the

207   corresponding conditions across the runs. Regressors of no interest were also included to

208   account for differences in the mean MR signal across scans and for head motion within scans.

209   In the univariate analysis, the regression weights (betas) from the GLM were compared

210   between conditions after averaging across the voxels of a region of interest (ROI). In the

211   multivariate analysis, the pattern of betas from the GLM across the voxels of an ROI were

212   compared between conditions using Kendall's tau correlation coefficient (τ) as a metric for

213   similarity. Before comparing the betas between the main and baseline experiments, the data

214   were mean-centered: the mean across all main experiment condition betas was subtracted from

215   those condition betas (separately for each voxel), and the mean across all baseline experiment

216   condition betas were subtracted from those condition betas.

217   *Regions of interest*

218   All ROIs were defined across both hemispheres (except FBA, which was limited to the right

219   hemisphere). In the multivariate analysis, we focused on two ROIs, the lateral-occipital cortex

220   (LOC) and the early visual cortex (EVC). The LOC ROI was defined using a group-constrained

221   subject-specific method (Fedorenko et al., 2010). The group-level ROI was defined by first

222   contrasting the average response to the 6 object categories with the response to the scrambled

223   objects in the baseline experiment. Threshold-free cluster enhancement (TFCE; Smith and

224   Nichols, 2009) with a permutation test was used to correct for multiple comparisons (at $p < 0.05$)

225   across the whole brain. The resulting voxels were intersected with the lateral occipital cortex

226   ROI from Julian et al. (Julian et al., 2012) to obtain the group-level LOC ROI. Then, for each

227   participant, the 1000 most object-selective voxels (average object response - scrambled

228   stimulus response, in the baseline experiment horizontal conditions) within the group-level LOC

229   ROI were selected for further analysis. The EVC ROI was defined at the individual participant

230   level as the 1000 most responsive voxels (average object response > 0, in the baseline

231   experiment horizontal conditions) in Brodmann area 17 (corresponding to V1; Wohlschläger et

232   al., 2005). Brodmann area 17 was taken from the Brodmann atlas available in SPM12.

233   In the univariate analysis we focused on two body-selective ROIs, the extrastriate body

234   area (EBA; Downing et al., 2001) and the fusiform body area (FBA; Peelen and Downing, 2005).

235   The ROIs were defined using the method described above for LOC. The group-level ROI was

236   defined by first contrasting the response to bodies with the average response to the other 5

237   categories in the baseline experiment. TFCE was used to correct for multiple comparisons (at p

238   < 0.05) across the whole brain. The resulting voxels were intersected with ROIs from Julian et

239   al. (2012): the extrastriate body area ROI to obtain the group-level EBA ROI and the fusiform

240   face area (FFA) ROI to obtain the group-level FBA ROI (FBA is not provided, but the FFA and

241   FBA closely overlap at the group-level; Peelen and Downing, 2005). Then, for each participant,

242   the 20 most body-selective voxels (body response - average response to other objects, in the

243   baseline experiment horizontal conditions) within the group-level ROIs were selected for further

244   analysis.

245   *Multivariate analysis approach*

246   In the multivariate analyses, we correlated multivoxel activity patterns evoked by the task-

247   irrelevant objects in the main experiment with multivoxel activity patterns evoked by the clearly

248   visible objects in the baseline experiment, using Kendall rank-ordered correlation; τ. We expect

249   to find stronger correlations between corresponding object categories (e.g., between bodies in

250   the main experiment and bodies in the baseline experiment), than between non-corresponding

251   categories (e.g. between bodies in the main experiment and beds in the baseline experiment).

252   As such, the difference between corresponding and non-corresponding category correlations is

253   a measure of category processing (Peelen et al., 2009), analogous to decoding accuracy. Here,

254   we computed *proximity* to the categories in the baseline experiment as the correlation with that

11

255   category minus the correlation with the other categories in the baseline experiment. For

256   example, for bodies, the proximity to bodies (in the baseline experiment) is the correlation

257   between bodies in the main experiment and bodies in the baseline experiment minus the

258   average correlation between bodies in the main experiment and the other five categories in the

259   baseline experiment.

260   *Image-based discriminability approach*

261   To rule out that bodies differed systematically from the other objects in terms of low-level

262   features, we used representations of the exemplars in the layers of a convolutional neural

263   network (trained for object recognition in natural images; CNN; AlexNet: Krizhevsky et al., 2012)

264   to test for image-based categorizability differences across the categories. Output activations at

265   each layer corresponding to 50 exemplars of each of the six categories, embedded in noise as

266   in the fMRI experiment, in the three possible locations defined by the shapes (see the

267   subsection on Stimuli), were extracted. Balanced linear support vector machines (SVM) were

268   trained to classify between the images of one category (150 images each) as opposed to the

269   other categories. 10-fold cross-validated classification accuracies were reported for each

270   category for each layer of the CNN.

271   **Results**

272   In the main experiment, participants detected the presence of object silhouettes belonging to

273   one of six categories (Fig. 1D), in different blocks. Throughout the experiment, only the

274   vertically-aligned locations were relevant for the detection task (Fig. 1A). Each block started with

275   a category cue (e.g. "Car") indicating the target category for that block (Fig. 1A), followed by 49

276   object detection trials. In 42 trials (6/7th), one of the two task-relevant locations contained a

277   briefly-presented object (67 ms) within phase-scrambled noise (Fig. 1B), with each category

278   presented equally often (7 trials each). In the remaining 7 trials (1/7th) no object was presented.

279   Crucially, in 6/7th of the trials, two objects were simultaneously presented in the

280   horizontally-aligned locations (Fig. 1A). These objects were briefly presented (67 ms),

281   embedded in noise, and backward masked. This was done to reduce the possibility of

282   participants moving their eyes (and/or spatial attention) to the task-irrelevant objects. Objects at

283   the horizontal locations were never relevant for the participants and participants were instructed

284   that these could be completely ignored. The occurrence probabilities of the categories were the

285   same as for the task-relevant locations. The 7 vertical and 7 horizontal conditions were fully

286   crossed within each block, resulting in 49 trials, which were presented in random order. Trials

287   were coded according to the categories presented in the horizontally aligned (task-irrelevant)

288   locations, as these were the focus of our analyses.

289   *Task performance (task-relevant locations)*

290   Averaged across the two fMRI sessions and across object search blocks, participants had a hit

291   rate of 78.3% (proportion of the target-present trials where participants responded) and a false

292   alarm rate of 5.6% (proportion of the target-absent trials where participants responded),

293   resulting in an average d' (zscore(hit rate) - zscore(false alarm rate)) of 2.7 (beds: 2.0; cars: 2.4;

294   bottles: 2.6; bodies: 2.9; chairs: 2.9; lamps: 3.3). Note that this was the performance for the

295   task-relevant stimuli presented at the vertical locations. All fMRI analyses focused on the

13

296   objects presented at the task-irrelevant horizontal locations. For responses to objects at the

297   task-irrelevant locations, we refer the reader to the section *The relationship between attentional*

298   *modulation and behavioral responses*.

299   *Univariate results in EBA and FBA*

300   Previous research has shown that bodies evoke a selective univariate response in two focal

301   regions of high-level visual cortex: the extrastriate body area (EBA; Downing et al., 2001) and

302   the fusiform body area (FBA; Peelen & Downing, 2005). Here, EBA and FBA were defined

303   based on responses in the baseline experiment (see Material and Methods). We tested for

304   spatially-global attention effects for bodies in these ROIs by comparing body-selective

305   responses in EBA and FBA evoked by task-irrelevant bodies across target detection blocks in

306   the main experiment. Betas were averaged across the voxels of each ROI to acquire one beta

307   per condition for each ROI. For each category, the beta corresponding to within-block trials in

308   which no objects were presented was subtracted to account for block effects. Responses to

309   non-body objects and non-body detection blocks were averaged, such that we had 4 values for

310   each ROI: body and non-body stimuli, presented in the body and non-body detection blocks.

311   The difference between body and non-body stimuli within each block is a measure of body
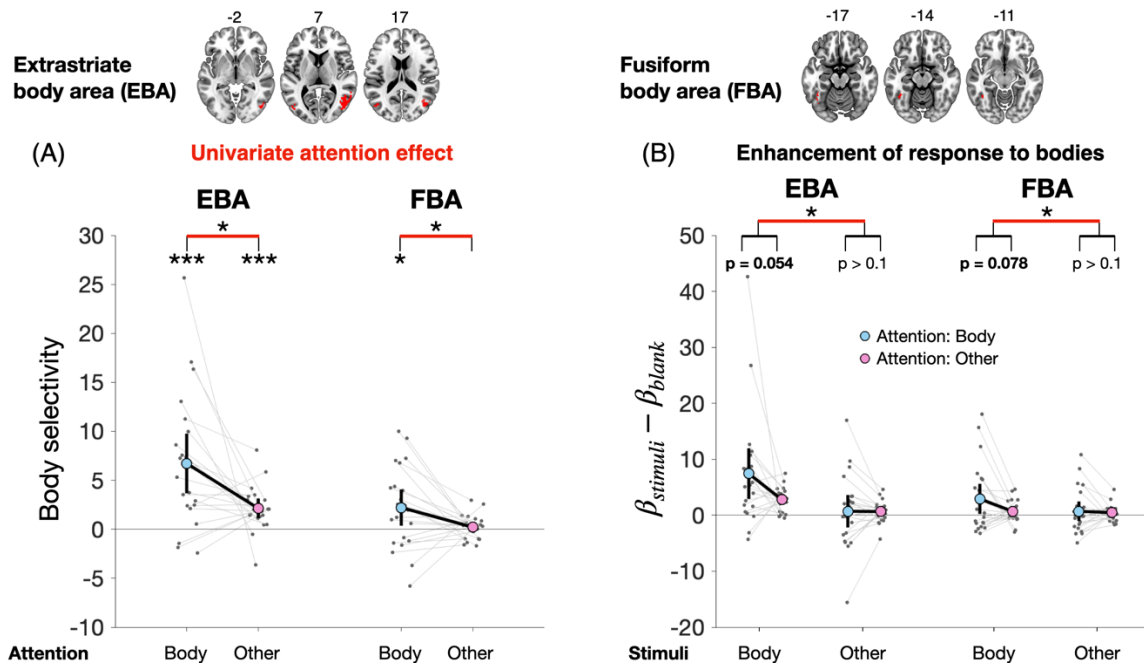
312   selectivity.

313   A 2 (ROI) x 2 (attention: body, other categories) repeated-measures ANOVA on body

314   selectivity (response to bodies minus average response to other categories) revealed a main

315   effect of attention ($F_{1,21}$ = 7.2, p = 0.014, η^2_p = 0.25), reflecting stronger body selectivity in body

316   attention blocks than non-body attention blocks (Fig. 2A). This attention effect interacted with

317   ROI ($F_{1,21}$ = 4.6, p = 0.043, η^2_p = 0.18), being stronger for EBA than FBA. When analyzed

318   separately, both EBA and FBA showed an attention effect, such that body selectivity was higher

319   in body detection blocks than in other category detection blocks (EBA: $F_{1,21}$ = 7.4, p = 0.013, η^2_p

320   = 0.26; FBA: $F_{1,21}$ = 4.4, p = 0.049, η^2_p = 0.17; Fig. 2A). The attention effect was consistent

14

321   across ROI sizes (Fig. 5). In EBA, bodies evoked a selective response in both the body

322   detection blocks (t_{21} = 4.6, p < 0.001, d = 1.00) and the other category detection blocks (t_{21} =

323   4.4, p < 0.001, d = 0.96), while in FBA body selectivity was only positive in the body detection

324   blocks (t_{21} = 2.5, p = 0.02, d = 0.55; other category detection blocks: t_{21} = 0.8, p = 0.42, d =

325   0.17).

326   The attention effect for bodies in EBA and FBA could reflect enhanced responses to

327   bodies presented in body detection blocks, but may also (or additionally) reflect reduced

328   responses (suppression) to the other categories presented in body detection blocks. To test

329   these alternatives, we compared body and object-evoked responses across the body and

330   object-detection blocks (after subtracting the response to blanks within each block). Averaged

331   across ROIs, there was a higher response to bodies in body detection blocks than in other

332   category detection blocks, which was marginally significant (paired t-test, t_{21} = 2.1, p = 0.05, d =

333   0.46; Fig. 2B). There was no evidence that the response to the other objects was suppressed,

334   with equally strong responses in both blocks (paired t-test, t_{21} = 0.19, p = 0.85, d = 0.04). These

335   effects were also observed, though weaker, in each ROI separately (statistics provided in Fig.

336   2B).

337   These results provide the first evidence for spatially-global attentional modulation for

338   body silhouettes, show that these effects are strongest in EBA, and link these effects to

339   enhancement of body responses rather than suppression of non-body responses.

340

15

Figure 2: Univariate attention effect in body-selective ROIs. (A) Body selectivity (response to body - average response to other objects) was higher when bodies were attended, in both ROIs. This provides evidence for spatially-global attentional modulation for body silhouettes. (B) Across ROIs, the response to bodies (corrected for block-wise differences by subtracting the corresponding blank responses) was enhanced while the responses to other categories remained unchanged. Error bars indicate 95% confidence intervals for the measures indicated on the y-axes. The asterisks indicate p-values for the t-tests of the corresponding comparisons (*p< 0.05, **p<0.01, ***p< 0.001). EBA and FBA ROIs are displayed together with MNI z-coordinates.

Multivariate results in LOC

Previous studies have shown that multivoxel activity patterns in object-selective cortex distinguish between object shapes (Haushofer et al., 2008; Op de Beeck et al., 2008; Eger et al., 2008). This gave us another opportunity to test for spatially-global effects of attention,

16

357   including for non-body categories. Here, instead of body selectivity, we used proximity (Pr) as

358   dependent measure. Proximity was based on correlations between response patterns in the

359   main experiment and response patterns in the baseline experiment, following previous work

360   (Peelen et al., 2009). Proximity reflects how similar a category's response pattern in the main

361   experiment is to a category's response pattern in the baseline experiment, relative to the other

362   categories in the baseline experiment (Materials and Methods). For example, for bodies, the

363   proximity to bodies (in the baseline experiment) is the correlation between bodies in the main

364   experiment and bodies in the baseline experiment minus the average correlation between

365   bodies in the main experiment and the other five categories in the baseline experiment.

366   *Attentional modulation for bodies in LOC*

367   The proximity to bodies is shown in Fig. 3A. A 2 (attention: body, other categories) x 2 (stimulus

368   presented: body, other categories) repeated-measures ANOVA revealed a significant interaction

369   ($F_{1,21}$ = 30.4, p < 0.001, η^2_p = 0.59), reflecting a stronger difference between the proximities for

370   body and non-body categories when participants attended to bodies (t_{21} = 9.9, p < 0.001, d =

371   2.2) than when they attended to the other categories (t_{21} = 8.1, p < 0.001, d = 1.8). The

372   multivariate attention effect in LOC was consistent across ROI sizes (Fig. 5). These results

373   provide further evidence for spatially-global attentional modulation of body processing.

374   The attention effect for bodies in LOC could reflect enhanced proximity to bodies for the

375   bodies presented in body detection blocks, but may also (or additionally) reflect reduced

376   proximity to bodies (suppression) for the other categories presented in body detection blocks.

377   To test for body-selective enhancement, we compared the proximity (to bodies in the baseline

378   experiment) for bodies in the body detection blocks with the corresponding proximity of other

379   objects in the body detection blocks. To account for overall differences between blocks (e.g.,

380   related to the cue or to block-based attentional bias), we subtracted the proximity to bodies for

381   the within-block trials in which no objects were presented. Results showed that proximity to

17

382   bodies was significantly enhanced for bodies presented in the body detection blocks as

383   compared with bodies presented in the other detection blocks (t_{21} = 3.5, p = 0.002, d = 0.76;

384   blue comparison in Fig. 3B). There was no evidence for suppression: proximity to bodies was

385   not different for objects presented in the body detection blocks as compared with objects

386   presented in the other detection blocks (t_{21} = 0.3, p = 0.78, d = 0.06; green comparison in Fig.

387   3B). The difference between these effects (red comparison in Fig. 3B) corresponds to the same

388   multivariate attention effect as shown in Fig. 3A. These results show that the multivariate

389   attention effect was primarily driven by the enhancement of body-selective response patterns, in

390   line with the univariate results (Fig. 2).

391
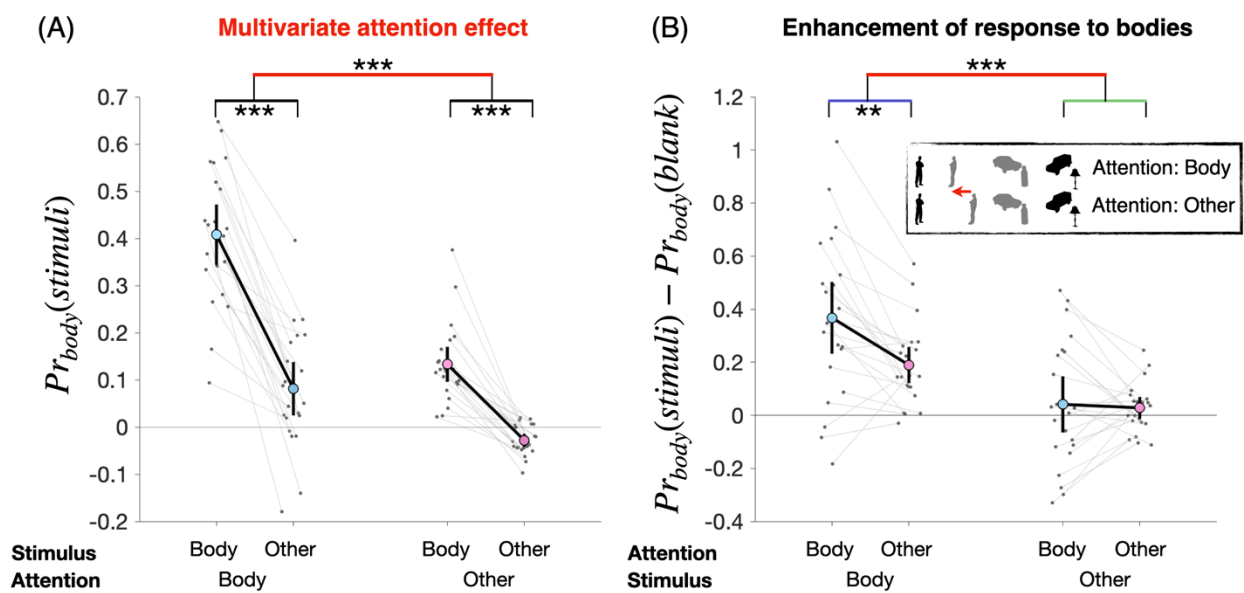
393   **Figure 3: Probing the multivariate attention effect for bodies in LOC.** (A) The selective proximity for

394   bodies (proximity to bodies for Body vs Other) is higher when bodies are attended, which is evidence for

395   a multivariate attention effect in LOC (comparison highlighted in red), reflecting response gain. (B)

396   Proximity (to bodies) of bodies and other categories were compared between the body attention blocks

397   and the other category attention blocks, corrected for block-wise differences by subtracting the proximity

398   (to bodies) to blank responses within blocks. When bodies were attended, the proximity of bodies was

399   enhanced, whereas the proximity of the other categories was not affected (inset: gray objects correspond

18

400   to attention-dependent representations and black to benchmark representations). This indicated that the

401   multivariate attention effect for bodies in LOC (the comparison corresponding to the red bar) was driven

402   primarily by enhancement of body-selective response patterns when bodies were attended. 95%

403   confidence intervals for the measures indicated on the y-axes are shown. The asterisks indicate the p-

404   values for the t-tests of the corresponding comparisons (*p< 0.05, **p<0.01, ***p< 0.001). Blue:

405   attentional modulation for bodies; green: attentional modulation for other categories.

406

407   *The relationship between attentional modulation and univariate body selectivity of LOC voxels*

408   Next, we tested whether the multivariate attention effect observed for bodies in LOC depended

409   on the (univariate) body-selectivity of voxels included in LOC. To this end, we computed the

410   multivariate attention effect for bodies in an ROI that consisted of LOC voxels that responded

411   less strongly to bodies than to other categories in the baseline experiment (on average 330.8

412   out of the original 1000 voxels satisfied this criterion). Results were compared with a size-

413   matched ROI consisting of randomly-sampled LOC voxels (size-matching done within each

414   participant; sampled 100 times). Attentional modulation was computed in the same way as for

415   the whole LOC in the original analysis (red comparison in Fig. 3). Attentional modulation was

416   stronger for the size-matched ROI than the non-selective ROI (t_{21} = 3.1, p = 0.006, d = 0.68).

417   However, attentional modulation was significant even in the non-selective ROI (t_{21} = 2.1, p =

418   0.047, d = 0.46). These results suggest that the attentional modulation in LOC was partly but

419   not exclusively driven by body-selective voxels.

420

421   *Attentional modulation for non-body categories in LOC*

422   Using the multivariate analysis framework outlined above for bodies, we can similarly test for

423   spatially-global attentional modulation for the other categories. For each non-body category, we

424   computed the multivariate attention effect as was done for bodies, now using the proximity to

19

425   that category in the baseline experiment. To reduce the complexity of the ANOVA and the

426   corresponding visualization of the data, we used *selective proximity* as the dependent measure.

427   Selective proximity is the proximity difference between the corresponding and non-

428   corresponding categories (e.g., the difference between the two left-most data points in Fig. 3A).

429   As an intuition for what this new measure represents, note that in the case of bodies, selective

430   proximity is analogous to the body selectivity measure in the univariate analysis.

431   In LOC, a 6 (category of interest) x 2 (category attended/unattended) repeated-

432   measures ANOVA on these selective proximities revealed a significant interaction ($F_{5,105}$ = 3.9,

433   p = 0.003, η^2_p = 0.16; Fig. 3A), indicating that attention differentially affected the selective

434   proximity of the six categories. Six paired-sample t-tests showed that attentional modulation was

435   significant for bodies (t_{21} = 5.5, p_{bonf} < 0.001, d = 1.2; red comparison in Fig. 4A), as already

436   shown in the previous analyses (Fig. 3). No significant multivariate attention effect was

437   observed for the other categories (t_{21} < 2.4, p_{bonf} > 0.1, d < 0.5; for all tests; Fig. 4A).

438   *Attentional modulation in EVC*

439   The same analysis was conducted in early visual cortex (EVC; see Materials and Methods). A 6

440   (category of interest) x 2 (category attended/unattended) repeated-measures ANOVA on

441   selective proximities revealed a marginally significant interaction ($F_{5,105}$ = 2.2, p = 0.06, η^2_p =

442   0.095; Fig. 4B), no significant main effect of attention ($F_{1,21}$ = 0.6, p = 0.4, η^2_p = 0.028), and a

443   marginally significant main effect of category ($F_{5,105}$ = 2.2, p = 0.06, η^2_p = 0.096). Paired-sample

444   t-tests showed no significant attentional modulation for any of the categories ($|t_{21}|$ < 2.2, p_{bonf} >

445   0.1, d < 0.48; for all tests). Finally, attentional modulation for bodies was significantly stronger in

446   LOC than in EVC (t_{21} = 2.9, p = 0.01, d = 0.63).

447
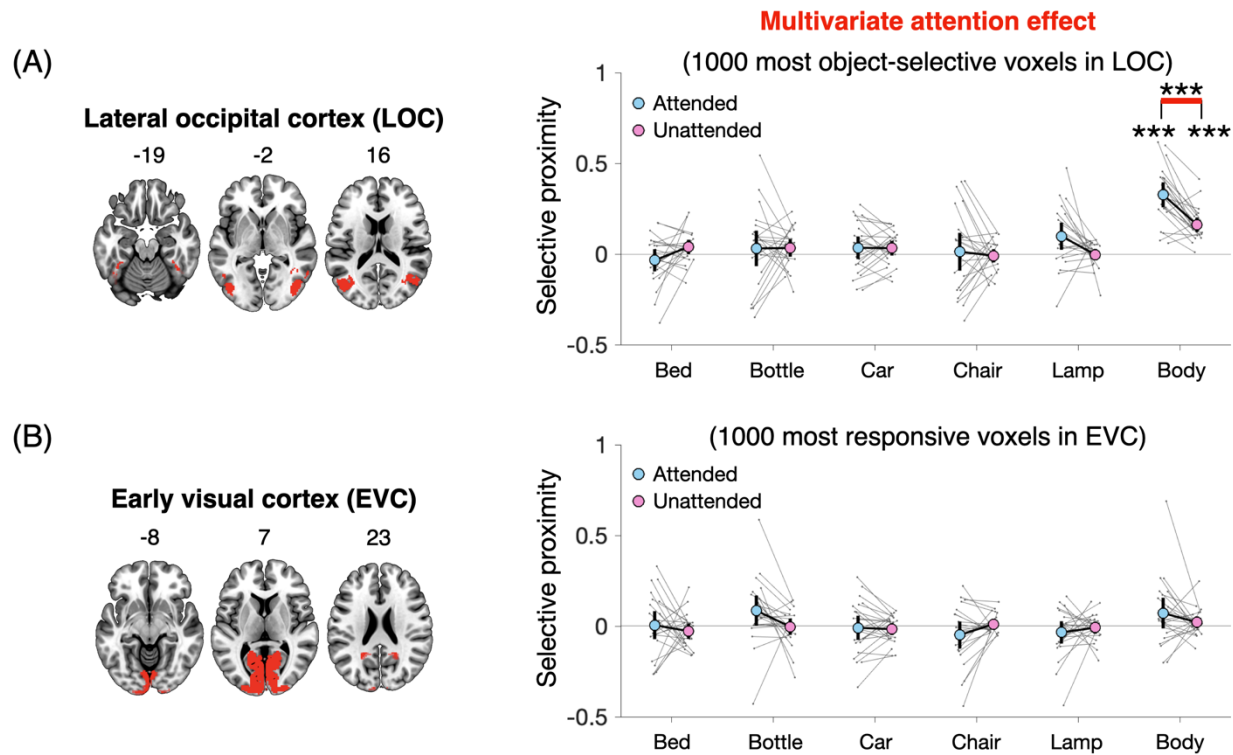
Figure 4: Multivariate attention effect. The selective proximities, for the attended and unattended conditions, are shown for the six categories in the two ROIs. The multivariate attention effect is the difference between attended and unattended selective proximity (comparison highlighted in red). A) In LOC, we find evidence for attentional modulation of the selective proximities of bodies. B) No attentional modulation was found in EVC. Error bars indicate 95% confidence intervals for the selective proximities. The asterisks denote Bonferroni corrected p-values for the t-tests of the twelve comparisons related to selective proximities, and Bonferroni corrected p-values for the t-tests of the six comparisons related to selective proximity modulations (*p< 0.05, **p<0.01, ***p< 0.001).

The relationship between attentional modulation and behavioral responses

In both multivariate and univariate analyses, we found that the body-selective response elicited by body silhouettes in task-irrelevant locations was enhanced in body detection blocks compared with other category detection blocks. This raises the question of whether this attentional modulation affected behavior in the detection task. Particularly, did participants

21

462   disproportionally false alarm to the bodies at task-irrelevant locations when detecting bodies at

463   task-relevant locations? Because of the orthogonal design, each category (+blank stimulus) in

464   the irrelevant location appeared equally often with each category (+blank stimulus) in the

465   relevant location. Therefore, when the target category (e.g., bodies) appeared at the task-

466   irrelevant location no target was presented at the task-relevant locations in most trials (6/7th),

467   and participants had to withhold their response. For these trials, we tested whether responses

468   (i.e., false alarms) depended on the combination of the category presented and the category

469   that was the target in that block. To this end, for each category, we computed the difference

470   between the false alarm rate (FA) to that category and the average FA to the other categories,

471   separately for each block. We then compared this ΔFA for trials in which the object matched the

472   target category (e.g., bodies presented in body blocks) and trials in which the object

473   mismatched the target category (e.g., bodies presented in bed blocks).

474   A 2 (matching, non-matching) x 6 (target category) repeated-measures ANOVA on ΔFA

475   revealed a significant interaction ($F_{5,105} = 3.3$, $p = 0.008$, $\eta^2_p = 0.14$; Fig. 6). Six paired-sample t-

476   tests showed that ΔFA was stronger when the object matched the target category for all

477   categories ($t_{21} > 2.9$, $p_{bonf} < 0.05$, $d > 0.63$, for all non-body categories, biggest difference of

478   6.5% for cars; bodies: $t_{21} = 2.79$, $p_{bonf} = 0.066$, difference of 3.7%, $d = 0.61$). These results show

479   that participants disproportionally false alarmed when the target category was shown at the

480   task-irrelevant location. Contrary to the fMRI results, however, this effect was relatively weak for

481   bodies.

482

22
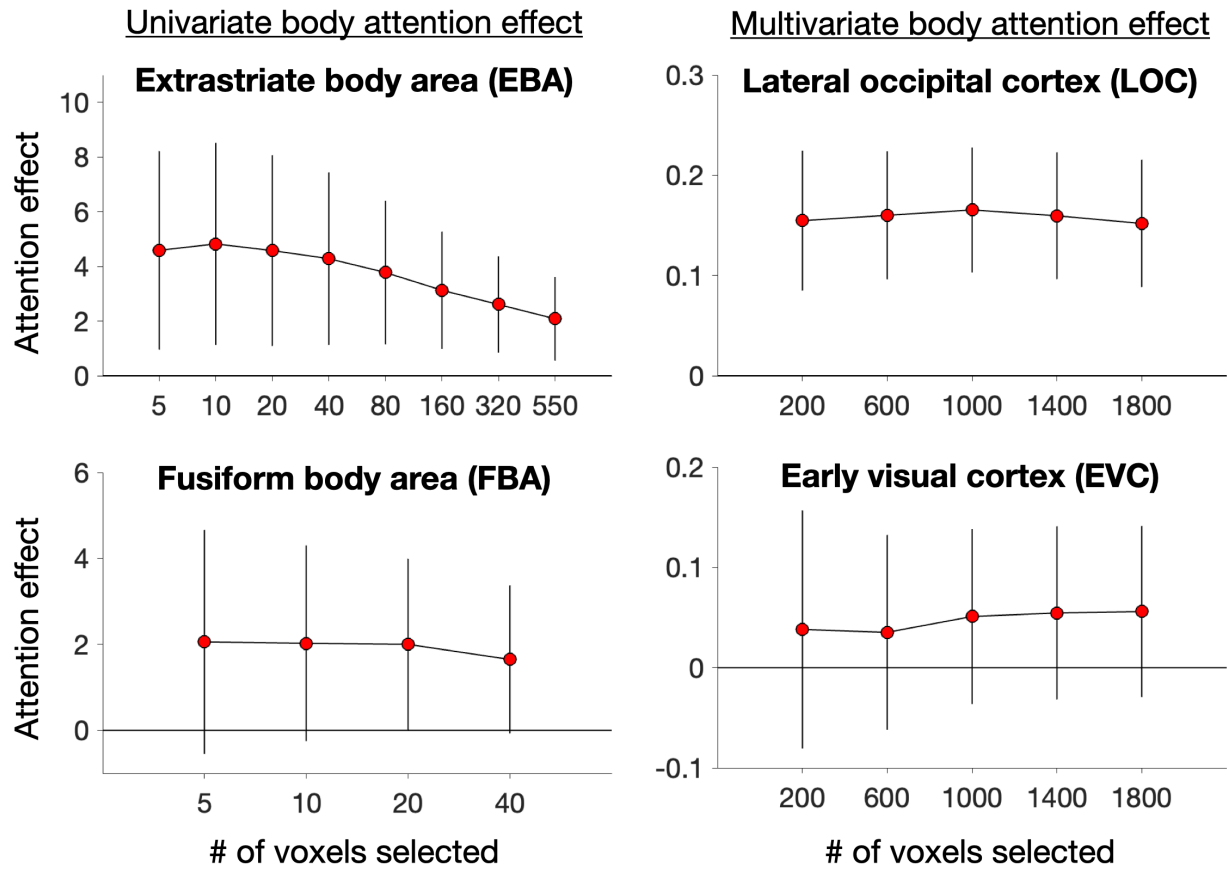
483

484

485   **Figure 5: The influence of voxel selection on body attention effects.** The univariate attention effects,

486   for EBA and FBA, and the multivariate attention effects, for LOC and EVC, for bodies, are shown as a

487   function of the number of voxels selected within each ROI. The attention effects for bodies observed in

488   EBA, FBA, and LOC, are observed regardless of the number of voxels selected. Error bars indicate 95%

489   confidence intervals for the attention effects.
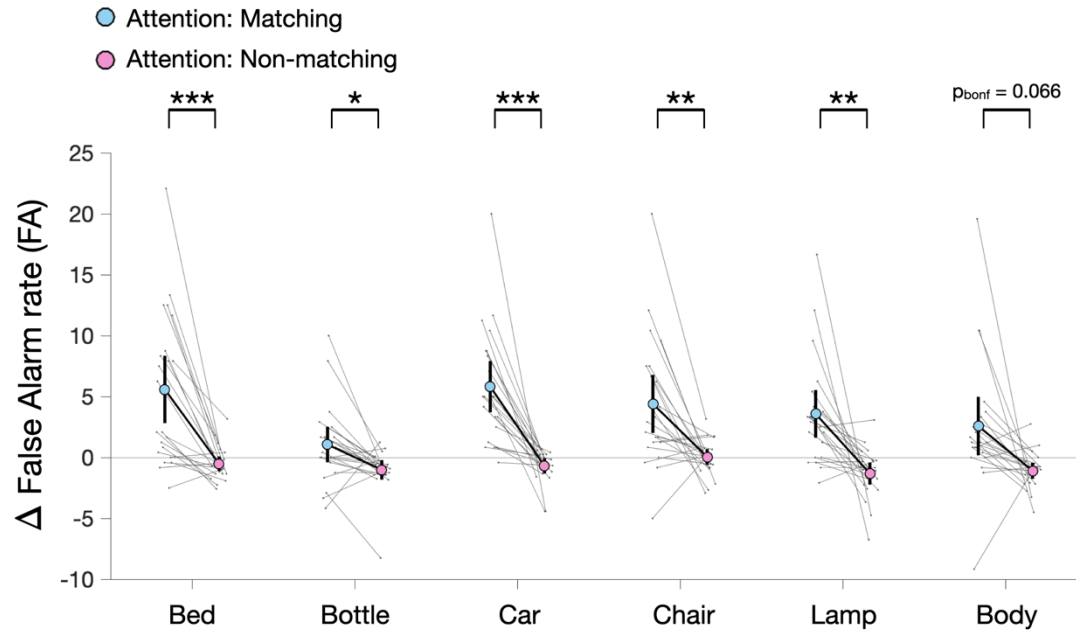
Figure 6: The relationship between attentional modulation and behavioral responses. Participants disproportionally false alarmed when the target category was shown at the task-irrelevant location (matching>non-matching) but this effect was relatively weak for bodies. Error bars indicate 95% confidence intervals for ΔFA. The asterisks denote instances where t-tests returned $p_{bonf} < 0.05$ for the corresponding comparisons (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).
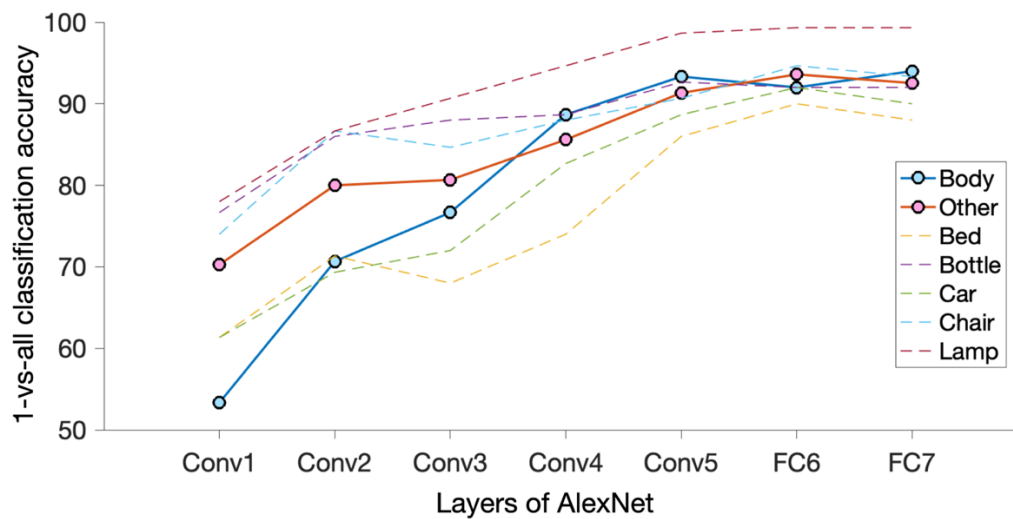
Image-based discriminability

In all fMRI analyses, we found that bodies were more strongly represented and more strongly modulated by attention than the other categories. This could reflect an interesting property of bodies, for example, related to the life-time relevance of detecting conspecifics or to the increased familiarity with body shapes. However, it could potentially also reflect uncontrolled image-based differences: perhaps the body silhouettes included in the study stood out from the other objects in terms of low-level features. To exclude this possibility, we decoded object categories from the object exemplar representations in the layers of a convolutional neural network trained for object recognition (Materials and Methods). For each of the 6 categories, in each layer of the CNN, one-vs-all linear discriminant classifiers were trained to discriminate

24

506   each category from the other categories using the 50 exemplars of each category presented in

507   the fMRI experiment. 10-fold cross-validation accuracies were analyzed across the objects.

508   As shown in Fig. 7, bodies were less discriminable than most other categories in the

509   early layers of the CNN. It is only in the mid to final layers - where overall classification is almost

510   at ceiling - that the classification accuracy for bodies is similar to the average accuracies for the

511   other categories. This result shows that the image-based discriminability was, if anything, lower

512   for bodies than for the other objects.

513

514

515   **Figure 7: Hierarchical image-based discriminability of the exemplars used in the fMRI experiment.**

516   One-vs-all classifiers were trained for each of the six categories, on the output activations of each layer of

517   a convolutional neural network trained for object recognition (AlexNet). 10-fold crossvalidation accuracies

518   are shown for all the objects in addition to the average accuracies for the non-body objects (termed

519   'Other'). Discriminability based on low-level features (corresponding to the early layers of the AlexNet)

520   was, if anything, lower for the human bodies than for the other objects. Therefore, it is unlikely that the

521   body-selective fMRI results reflect a distinct low-level property of bodies. 'Conv' refers to the convolutional

522   layers of AlexNet and 'FC' refers to the fully-connected layers.

Discussion

Across multiple analyses, we found convincing evidence that attention to human bodies enhanced visual cortex responses selective to bodies presented at task-irrelevant locations. This modulation reflected response gain rather than a generic bias, and could not be explained by low-level feature similarity of bodies. These results indicate that spatially-global attentional modulation – a hallmark of feature-based attention – can be found for features diagnostic of the presence of the human body.

The attentional effects observed here for body silhouettes are unlikely to reflect attention to low-level features such as orientation or color, for several reasons. First, we included a relatively large number of object categories in the experiment to ensure that participants could not detect objects based on low-level features, as these were shared with other categories (e.g., bottles were vertical, similar to bodies). Second, we presented object silhouettes instead of photographs to avoid possible low-level differences between categories in texture or color. Third, the image-based discriminability for each category, established using a convolutional neural network (CNN), indicated that bodies were difficult to discriminate from other categories based on low-level features encoded in the early layers of the CNN. Finally, the fMRI results showed attentional modulation in object-selective cortex (LOC) and body-selective EBA/FBA, but not early visual cortex (EVC), indicating an attentional modulation at a higher level of visual processing.

Our results are in line with the feature similarity gain modulation model (FSGM; Maunsell & Treue, 2006) by showing that feature-based attention enhanced the response to the voxels' preferred stimuli. Specifically, attention to bodies made the response pattern evoked by task-irrelevant bodies more similar to prototypical body response patterns established in a separate baseline experiment. Furthermore, these attention effects were strongest in body-selective voxels of LOC. Finally, reliable univariate attention effects were observed in independently-

26

548   defined body-selective regions (EBA/FBA). It should be noted that we did not find evidence that

549   responses to the other categories were suppressed, as proposed by FSGM. However, the

550   response to other categories was low and any suppression (posited to be smaller in magnitude

551   than enhancement by FSGM) might not be observable in this case.

552   The finding of spatially-global modulation for human bodies adds to previous evidence

553   for global modulation for faces. Specifically, in one study, peripherally presented and task-

554   irrelevant faces evoked a stronger face-selective N170 electro-encephalography (EEG)

555   response when participants attended to faces than to houses (Störmer et al., 2019).

556   Furthermore, in fMRI, responses to peripheral faces in the face-selective fusiform face area

557   (FFA) were more strongly modulated by the task-set of the participants (i.e., whether or not they

558   focused on faces) than by spatial attention (Reddy et al., 2007). Together with the current

559   findings, these results provide evidence for spatially-global attentional modulation for bodies and

560   faces, two socially relevant categories that are selectively represented in the visual cortex

561   (Downing et al., 2006; Kanwisher, 2010).

562   While these results suggest that bodies and faces may be special – reflecting their

563   unique social and biological significance – we do not rule out that spatially-global attentional

564   modulation may also exist for other highly-familiar object categories. For example, behavioral

565   studies showed that animals and vehicles could be detected in the near-absence of spatial

566   attention (Li et al., 2002; but see Cohen et al., 2011), with category-based attention facilitating

567   object detection independently of spatial attention (Stein and Peelen, 2017). Indeed, based on

568   the overlap in human and animal features in detection tasks (Evans and Treisman, 2005), it is

569   plausible that our results would generalize to other animals, particularly those that activate

570   body-selective regions (Downing et al., 2006). Similarly, extensive experience with particular

571   objects may drive selective neural tuning (Gauthier and Logothetis, 2000; McGugin et al., 2012;

572   Frank et al., 2014) and give rise to similar behavioral advantages as those observed for bodies

573   (Hershler and Hochstein, 2009; Golan et al., 2014; Reeder et al., 2016; Stein et al., 2016).

574   Taking everything together, the evidence suggests that features that are diagnostic of

575   bodies meet many of the previously proposed criteria for basic features: showing spatially-global

576   attentional modulation (Maunsell and Treue, 2006), being processed "early, automatically, and

577   in parallel across the visual field" (Treisman and Gelade, 1980), and being represented

578   selectively in the visual system (Treisman, 2006). Indeed, Treisman (2006) proposed that the

579   feature detectors of the feature integration theory are not necessarily limited to low-level

580   features such as orientation and color. Raising the possibility that there may be animal feature

581   detectors, Treisman (2006) noted that animal features may not necessarily be more complex for

582   the visual system than colors, line orientations, or direction of motion. By providing evidence for

583   spatially-global attentional modulation for human bodies, our results support this proposal.

584   Our findings raise the question of what features are attended when attention is directed

585   to bodies. Addressing this question for animals, Treisman (2006) suggested that: "participants

586   may be set to sense, in parallel, a highly overlearned vocabulary of features that characterize a

587   particular semantic category." One possibility is thus that attention to bodies is mediated by

588   attention to a set of mid-level features that are diagnostic of human bodies (Ullman et al., 2002;

589   Reeder and Peelen, 2013). Alternatively, attention may be directed to holistic representations of

590   body shape (Reed et al., 2003; Stein et al., 2012). Future studies could test these alternatives

591   by measuring global attentional modulation for various body-related features, body parts, and

592   inverted bodies at the task-irrelevant location while participants attend to bodies at the task-

593   relevant locations (Reeder and Peelen, 2013).

594   To conclude, the current results provide evidence for spatially-global attentional

595   modulation for human bodies in high-level visual cortex, linking this modulation to body-selective

596   representations in univariate and multivariate analyses. Combining these results with previous

597   behavioral and neuroimaging studies, we propose that bodies may be processed as basic

598   features, supporting the rapid and parallel detection of conspecifics in our environment even

599   outside the focus of spatial attention.

28

600 **References**

601   Carrasco M (2011) Visual attention: The past 25 years. Vision Research 51:1484–1525.

602   Chun MM, Golomb JD, Turk-Browne NB (2011) A Taxonomy of External and Internal Attention.

603   Annu Rev Psychol 62:73–101.

604   Cohen MA, Alvarez GA, Nakayama K (2011) Natural-Scene Perception Requires Attention.

605   Psychol Sci 22:1165–1172.

606   Downing PE, Bray D, Rogers J, Childs C (2004) Bodies capture attention when nothing is

607   expected. Cognition 93:B27–B38.

608   Downing PE, Chan AW-Y, Peelen MV, Dodds CM, Kanwisher N (2006) Domain Specificity in

609   Visual Cortex. Cerebral Cortex 16:1453–1461.

610   Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A Cortical Area Selective for Visual

611   Processing of the Human Body. Science 293:2470–2473.

612   Evans KK, Treisman A (2005) Perception of Objects in Natural Scenes: Is It Really Attention

613   Free? Journal of Experimental Psychology: Human Perception and Performance 31:1476–1492.

614   Fedorenko E, Hsieh P-J, Nieto-Castañón A, Whitfield-Gabrieli S, Kanwisher N (2010) New

615   Method for fMRI Investigations of Language: Defining ROIs Functionally in Individual Subjects.

616   Journal of Neurophysiology 104:1177–1194.

617   Frank SM, Reavis EA, Tse PU, Greenlee MW (2014) Neural mechanisms of feature conjunction

618   learning: Enduring changes in occipital cortex after a week of training: Neural Mechanisms of

619   Feature Conjunction Learning. Hum Brain Mapp 35:1201–1211.

620   Gauthier I, Logothetis NK (2000) Is face recognition not so unique after all? Cognitive

621   Neuropsychology 17:125–142.

622   Golan T, Bentin S, DeGutis JM, Robertson LC, Harel A (2014) Association and dissociation

623   between detection and discrimination of objects of expertise: Evidence from visual search. Atten

624   Percept Psychophys 76:391–406.

625 Haushofer J, Livingstone MS, Kanwisher N (2008). Multivariate patterns in object-selective

626 cortex dissociate perceptual and physical shape similarity. PLoS Biol. 6 (7), e187.

627 Hershler O, Hochstein S (2009) The importance of being expert: Top-down attentional control in

628 visual search with photographs. Attention, Perception & Psychophysics 71:1478–1486.

629 Jehee JFM, Brady DK, Tong F (2011) Attention Improves Encoding of Task-Relevant Features

630 in the Human Visual Cortex. Journal of Neuroscience 31:8210–8219.

631 Julian JB, Fedorenko E, Webster J, Kanwisher N (2012) An algorithmic method for functionally

632 defining regions of interest in the ventral visual pathway. NeuroImage 60:2357–2364.

633 Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. Nat

634 Neurosci 8:679–685.

635 Kanwisher N (2010) Functional specificity in the human brain: A window into the functional

636 architecture of the mind. Proceedings of the National Academy of Sciences 107:11163–11170.

637 Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional

638 neural networks. Adv Neural Inform Process Syst 1:1097–1105.

639 Li FF, VanRullen R, Koch C, Perona P (2002) Rapid natural scene categorization in the near

640 absence of attention. Proceedings of the National Academy of Sciences 99:9596–9601.

641 Liu T, Larsson J, Carrasco M (2007) Feature-Based Attention Modulates Orientation-Selective

642 Responses in Human Visual Cortex. Neuron 55:313–323.

643 Maunsell JHR, Treue S (2006) Feature-based attention in visual cortex. Trends in

644 Neurosciences 29:317–322.

645 McGugin RW, Gatenby JC, Gore JC, Gauthier I (2012) High-resolution imaging of expertise

646 reveals reliable object selectivity in the fusiform face area related to perceptual performance.

647 Proceedings of the National Academy of Sciences 109:17063–17068.

648 Op de Beeck HP, Torfs K, Wagemans J (2008). Perceived shape similarity among unfamiliar

649 objects and the organization of the human object vision pathway. J. Neurosci. 28 (40), 10111–

650 10123.

651   Peelen MV, Downing PE (2005) Selectivity for the Human Body in the Fusiform Gyrus. Journal

652   of Neurophysiology 93:603–608.

653   Peelen MV, Fei-Fei L, Kastner S (2009) Neural mechanisms of rapid natural scene

654   categorization in human visual cortex. Nature 460:94–97.

655   Peelen MV, Kastner S (2011) A neural basis for real-world visual search in human

656   occipitotemporal cortex. Proceedings of the National Academy of Sciences 108:12125–12130.

657   Reddy L, Moradi F, Koch C (2007) Top–down biases win against focal attention in the fusiform

658   face area. NeuroImage 38:730–739.

659   Reed CL, Stone VE, Bozova S, Tanaka J (2003) The body-inversion effect. Psychological

660   Science 14:302–308.

661   Reeder RR, Peelen MV (2013) The contents of the search template for category-level search in

662   natural scenes. Journal of Vision 13:13–13.

663   Reeder RR, Stein T, Peelen MV (2016) Perceptual expertise improves category detection in

664   natural scenes. Psychon Bull Rev 23:172–179.

665   Reeder RR, van Zoest W, Peelen MV (2015) Involuntary attentional capture by task-irrelevant

666   objects that match the search template for category detection in natural scenes. Atten Percept

667   Psychophys 77:1070–1080.

668   Ro T, Friggel A, Lavie N (2007) Attentional biases for faces and body parts. Visual Cognition

669   15:322–348.

670   Saenz M, Buracas GT, Boynton GM (2002) Global effects of feature-based attention in human

671   visual cortex. Nat Neurosci 5:631–632.

672   Serences JT, Boynton GM (2007) Feature-Based Attentional Modulations in the Absence of

673   Direct Visual Stimulation. Neuron 55:301–312.

674   Smith S, Nichols T (2009) Threshold-free cluster enhancement: Addressing problems of

675   smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44:83–98.

676 Stein T, Peelen MV (2017) Object detection in natural scenes: Independent effects of spatial

677 and category-based attention. Atten Percept Psychophys 79:738–752.

678 Stein T, Reeder RR, Peelen MV (2016) Privileged Access to Awareness for Faces and Objects

679 of Expertise. Journal of Experimental Psychology: Human Perception and Performance 42:788–

680 798.

681 Stein T, Sterzer P, Peelen MV (2012) Privileged detection of conspecifics: Evidence from

682 inversion effects during continuous flash suppression. Cognition 125:64–79.

683 Störmer VS, Cohen MA, Alvarez GA (2019) Tuning Attention to Object Categories: Spatially

684 Global Effects of Attention to Faces in Visual Processing. Journal of Cognitive Neuroscience

685 31:937–947.

686 Treisman A (2006) How the deployment of attention determines what we see. Visual Cognition

687 14:411–443.

688 Treisman AM, Gelade G (1980) A Feature-Integration Theory of Attention. Cognitive Psychology

689 12:97–136.

690 Treue S, Trujillo JCM (1999) Feature-based attention influences motion processing gain in

691 macaque visual cortex. Nature 399:575–579.

692 Ullman S, Vidal-Naquet M, Sali E (2002) Visual features of intermediate complexity and their

693 use in classification. Nat Neurosci 5:682–687.

694 Wohlschläger AM, Specht K, Lie C, Mohlberg H, Wohlschläger A, Bente K, Pietrzyk U, Stöcker

695 T, Zilles K, Amunts K, Fink GR (2005) Linking retinotopic fMRI mapping and anatomical

696 probability maps of human occipital areas V1 and V2. NeuroImage 26:73–82.

697 Wolfe JM (1994) Guided search 2.0. A revised model of visual search. Psychon Bull Rev 1:202–

698 238.

699 Xiao J, Hays J, Ehinger KA, Oliva A, Torralba A (2010) SUN database: Large-scale scene

700 recognition from abbey to zoo. In: 2010 IEEE Computer Society Conference on Computer

701   Vision and Pattern Recognition, pp 3485–3492. San Francisco, CA, USA: IEEE. Available at:

702   http://ieeexplore.ieee.org/document/5539970/ [Accessed October 1, 2021].

703   Zhang W, Luck SJ (2009) Feature-based attention modulates feedforward visual processing.

704   Nat Neurosci 12:24–25.

705