1   **Typical retinotopic locations impact the time course of object coding**

2

3   Daniel Kaiser[1,*], Merle M. Moeskops[1], Radoslaw M. Cichy[1,2,3]

4

5   *[1]Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany*

6   *[2]Berlin School of Mind and Brain, Humboldt-Universität Berlin, Berlin, Germany*

7   *[3]Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany*

8

9   *[*]Correspondence to:*

10   Dr. Daniel Kaiser

11   Department of Education and Psychology

12   Freie Universität Berlin

13   Habelschwerdter Allee 45

14   14195 Berlin, Germany

15   danielkaiser.net@gmail.com

16   **Abstract**

17   In everyday visual environments, objects are non-uniformly distributed across visual

18   space. Many objects preferentially occupy particular retinotopic locations: for example,

19   lamps more often fall into the upper visual field, whereas carpets more often fall into the

20   lower visual field. The long-term experience with natural environments prompts the

21   hypothesis that the visual system is tuned to such retinotopic object locations. A key

22   prediction is that typically positioned objects should be coded more efficiently. To test

23   this prediction, we recorded electroencephalography (EEG) while participants viewed

24   briefly presented objects appearing in their typical locations (e.g., an airplane in the upper

25   visual field) or in atypical locations (e.g., an airplane in the lower visual field). Multivariate

26   pattern analysis applied to the EEG data revealed that object classification depended on

27   positional regularities: Objects were classified more accurately when positioned typically,

28   rather than atypically, already at 140 ms, suggesting that relatively early stages of object

29   processing are tuned to typical retinotopic locations. Our results confirm the prediction

30   that long-term experience with objects occurring at specific locations leads to enhanced

31   perceptual processing when these objects appear in their typical locations. This may

32   indicate a neural mechanism for efficient natural scene processing, where a large number

33   of typically positioned objects needs to be processed.

34

35   **Keywords**

36   visual perception, object recognition, real-world regularities, location priors, EEG

37   decoding, multivariate pattern analysis

1 Introduction

Visual objects are enclosed entities that can in principle be moved around freely. However, in everyday environments object positions are often quite constrained. For instance, consider the predictability in the locations of objects in a living room: The sofa is facing the TV, a table is in between the two, a lamp hangs from the ceiling, whereas carpets lie on the floor. This example illustrates that the object content of natural scenes is organized in repeatedly occurring positional structures (Bar, 2000; Chun, 2002). Many previous studies have investigated how inter-object relationships in these positional structures (e.g., lamps appearing above tables) impact behavioral performance and neural processing (Biederman, Mezzanotte, & Rabinowitz, 1982; Kaiser, Stein, & Peelen, 2014; Oliva & Torralba, 2007; Wolfe, Võ, Evans, & Greene, 2011). However, positional object structures often also imply that individual objects are associated with particular locations in space (e.g., lamps appearing in the upper part of a scene). It has recently been proposed that the visual system is tuned to these regularities (Kaiser & Haselhuhn, 2017; Kravitz, Vinson, & Baker, 2008), which could facilitate neural processing for individual objects appearing in retinotopic locations that correspond to their typical real-world locations.

Such location-specific variations in object coding are suggested by previous results that indicate the co-representation of object identity and location information in visual cortex: (1) cortical responses depend on the position of the object in the visual field (Hemond, Kanwisher, & Op de Beeck, 2007; Hasson, Levy, Behrmann, Hendler, & Malach, 2002), (2) object selective cortex contains information about both an object's identity and its location (Cichy, Chen, & Haynes, 2011; Golomb & Kanwisher, 2011; Hong, Yamins, Majaj, & DiCarlo, 2017; Kravitz, Kriegeskorte, & Baker, 2010; Schwarzlose, Swisher, Dang, &

62   Kanwisher, 2008), and (3) information about object identity and location emerge at

63   similar time points in visual processing (Isik, Meyers, Leibo, & Poggio, 2014; Carlson,

64   Hogendoorn, Kanai, Mesik, & Turret, 2011).

65   The link between identity and location information in object processing creates

66   the possibility that the two properties interact. In everyday environments, the visual

67   system is repeatedly faced with positional structures, where individual object positions

68   are highly predictable. Under typical viewing conditions, and unless directly fixated,

69   objects appearing in the upper part of scenes also more often occupy locations in the

70   upper visual field, while objects appearing in the lower part of scenes are repeatedly

71   encountered in the lower visual field. Through repeated exposure, retinotopic object-

72   coding mechanisms could get tuned to typical object locations in the visual field. Thus,

73   over time, neural channels are shaped to represent typical object-location conjunctions in

74   a maximally efficient way. These efficient location-specific object representations would

75   enhance the processing of an object when it appears in its typical locations within a scene

76   – and within the visual field. Evidence for such a processing enhancement has been found

77   in the domain of person perception, where typical configurations impact cortical

78   responses to individual face and body parts (Chan, Kravitz, Truong, Arizpe, & Baker, 2010;

79   de Haas et al., 2016; Henriksson, Mur, & Kriegeskorte, 2015). For example, in face-selective

80   visual cortex, response patterns are better discriminable for typically, as compared to

81   atypically, positioned face parts (de Haas et al., 2016), revealing visual processing

82   channels that are tuned to the spatial regularities in the face.

83   Here, we test the prediction that the positional regularities contained in natural

84   scenes can similarly facilitate the processing of everyday objects appearing in their typical

85   retinotopic locations. Participants viewed objects associated with upper and lower visual

86   field locations (e.g., a lamp or a carpet) (Figure 1A) while we recorded

87   electroencephalography (EEG). We used multivariate classification on the EEG data

88   (Contini, Wardle, & Carlson, 2017) to track the time course of object coding with high

89   temporal precision. Analyses revealed that after 140ms the visual processing of an object

90   is affected by its typical location in the visual field: Objects appearing in their typical

91   locations (e.g., a lamp in the upper visual field and a carpet in the lower visual field) could

92   be decoded more successfully than objects appearing in atypical locations (e.g., a carpet

93   in the upper visual field and a lamp in the lower visual field). These results suggest that

94   early stages of visual processing are tuned to the positional object structure of real-world

95   scenes.

2 Materials and Methods

2.1 Participants

Thirty-four healthy adults (mean age 26.4 years, SD = 5.4; 23 female) completed the experiment. The sample size was set a-priori, based on considerations regarding statistical power: A sample size of 34 is needed for detecting a simple effect with a medium effect size of d = 0.5 with a probability of more than 80%. All participants had normal or corrected-to-normal vision, provided informed consent and received monetary reimbursement or course credits for their participation. All procedures were approved by the ethical committee of the Department of Education and Psychology of the Freie Universität Berlin and were in accordance with the Declaration of Helsinki.

2.2 Stimuli

The stimulus set consisted of greyscale images of six objects associated with typical visual field locations, of which three were associated with upper visual field locations (lamp, airplane, and hat) and three were associated with lower visual field locations (carpet, boat, and shoe). For each object, ten exemplars were used (see Figure 1A for stimulus examples). To avoid a confounding of visual field associations and other, conceptual stimulus attributes, the objects were matched for their categorical content (two furniture items, two transportation items, and two clothing items), so that high-level properties like the objects' size, manipulability and semantic associations were comparable across upper and lower visual field objects. To control for low-level visual factors, all stimulus images were gray-scaled and matched for overall luminance (using the SHINE toolbox; Willenbockel et al., 2010). Additionally, as a measure of low-level image similarity, pixel-wise similarity was computed between all 60 images (i.e., 6 objects, 10 exemplars each), in a pairwise fashion: Pixel intensity values were correlated between all same visual field

120 objects (e.g., lamps and airplanes) and between all different visual field objects (e.g.,

121 lamps and boats). No difference between all possible (600) within-field and all possible

122 (900) between-field correlations emerged, $t(1498)=0.50$, $p=0.62$, indicating that objects

123 with identical typical locations were no more similar in their low-level visual properties

124 than objects with different typical locations.

125 To ensure that the six objects could be reliably linked to a specific location, we

126 validated the association of the six objects with a specific part of the visual field in two

127 ways. First, we assessed the typical spatial distribution of each object in natural scenes,

128 assuming that natural scene photographs represent a snapshot of the visual field roughly

129 approximating natural viewing conditions. Hence, the distribution of the objects in the

130 scene image should be similar to their distribution across the visual field. To objectively

131 measure the typical position of each object within a scene, we queried a huge number of

132 scene photographs (>100,000) from the LabelMe toolbox, where human observers

133 annotated single objects by drawing labelled polygons (Russell, Torralba, Murphy, &

134 Freeman, 2008). For all scenes containing a specific object we computed the mean pixel

135 coordinate of the area labeled as belonging to the object and then averaged these

136 positions across scenes. The resulting "typical" object locations showed that, as

137 expected, the upper visual field objects were associated with locations (y: vertical

138 coordinate from bottom (0) to top (1) of the scene) in the upper parts of scenes (lamp: y

139 = 0.61, SD = 0.17; airplane: y = 0.52, SD = 0.20; hat: y = 0.60, SD = 0.18), while lower visual

140 field objects were associated with locations in the lower part of scenes (carpet: y = 0.17,

141 SD = 0.13; boat: y = 0.42, SD = 0.18; shoe: y = 0.37, SD = 0.17). The typical location in scenes

142 differed significantly between objects associated with the upper and lower visual field, t >

143 11.4, $p < .001$, for all pairwise comparisons. Figure 1B shows the distribution of object

144 locations along the vertical axis of the scenes, split into 7 bins.

145 Second, we sought to demonstrate a correspondence between this automated,

146 scene-based measure and people's explicit associations of the objects with particular

147 locations in space. We thus asked a set of participants (n = 70 for lamp/carpet, n = 60 for

148 airplane/boat, n = 65 for hat/shoe; including the participants of the current study, after

149 the completion of the EEG experiment) to indicate the typical locations in which they

150 expect to see each of the six objects. In this task, participants were asked to drag the

151 image of a single exemplar of each object to its typical location on a computer screen

152 (imagining that the computer screen represents their field of view in a natural scene). The

153 central part of the screen – where the object initially appeared – was blocked (indicated

154 by a grey circle), so that participants (1) could not place the object in a central location of

155 the screen, and (2) had to move the object before proceeding. As expected, participants

156 more often chose upper screen positions (y: vertical coordinate from bottom (0) to top

157 (1) of the screen) for the upper visual field objects (lamp: $y = 0.65$, $SD = 0.19$; airplane: $y =$

158 0.67, $SD = 0.20$; hat: $y = 0.57$, $SD = 0.20$), and lower screen positions for the lower visual

159 field objects (carpet: $y = 0.29$, $SD = 0.22$; boat: $y = 0.36$, $SD = 0.18$; shoe: $y = 0.30$, $SD = 0.18$).

160 The vertical locations chosen by the participants differed significantly between objects

161 associated with the upper and lower visual field, $t > 6.04$, $p < .001$, for all pairwise

162 comparisons. Figure 1C shows the distribution of vertical object locations on the screen,

163 split into 7 bins. The scene-based measure and participants' explicit assessment thus

164 provided converging evidence for the association of the objects with specific spatial
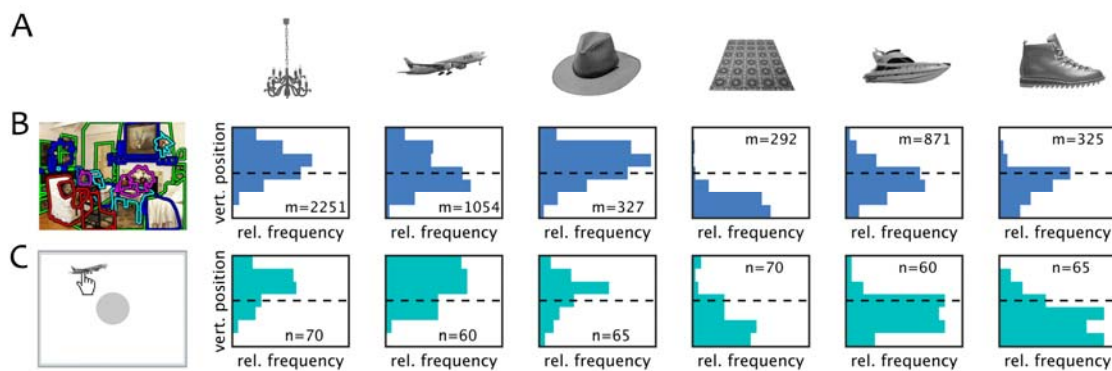
165 locations.

166

Figure 1. A) Example Stimuli. Ten different exemplar images of six objects each (here, one example image for each object shown) were used as stimuli, of which three objects were associated with the upper visual field (lamp, airplane, hat) and three were associated with the lower visual field (carpet, boat, shoe). B) To validate our assessment of visual field associations, we automatically extracted the positions for each object in a large set of labelled scene photographs taken from the LabelMe scene database (Russell et al., 2008). For each scene, we determined the relative position of the object along the vertical axis, and plotted the distribution across 7 bins (*m*: number of scenes for each object). C) Additionally, we asked a group of participants to indicate for each object the position that it typically occupies in the visual field by dragging the object to the desired location. We then computed the distribution of relative locations along the vertical axis of the screen, split into 7 bins (*n*: number of participants that indicated the typical location for each object). Both measures confirmed the spatial priors associated with the six objects.

2.3 Experimental Design

To test whether objects are processed differently when presented in typical and atypical locations within the visual field, the objects were presented in the upper or lower visual field (Figure 2A). On every trial, one object exemplar was presented in one of the

187   two locations for 150ms, followed by a variable inter-trial interval (randomly jittered, from

188   1250ms to 1750ms). The brief presentation time was chosen to not give participants

189   enough time to move their eyes towards the objects. Stimuli were presented at 3.25°

190   vertical eccentricity and subtended a visual angle of maximally 3° in horizontal and vertical

191   axes. Stimulus presentation was controlled using the Psychtoolbox (Brainard, 1997).

192   Participants were asked to detect one-back repetitions on an object level (e.g., two

193   different airplanes in direct succession; see Figure 2A). Repetitions occurred on 13% of the

194   trials and equally often for typically and atypically positioned repetition targets and for

195   the top and bottom locations. Participants performed accurately on this task (mean

196   accuracy 96%, $SD = 2\%$), with no difference in accuracy between typically and atypically

197   positioned objects, $t(33) = 0.87$, $p = .391$. One-back repetition trials (i.e., the "second" trial

198   containing the repetition) were removed from all EEG analyses. The whole experiment

199   consisted of 1656 trials: in 1440 trials, each object exemplar was shown 12 times in each

200   location (i.e., 120 repetitions per object and location), while the remaining 216 trials were

201   one-back repetition trials. The experiment was split into 8 runs, and participants could

202   take breaks between the runs. Twelve participants completed an extended experimental

203   session with 2760 trials (including 360 repetition trials), which additionally contained the

204   same conditions at large eccentricities in half of the trials; these additional data are not

205   reported here. The 2400 non-repetition trials consisted of 10 repetitions of each object

206   exemplar in each of the four locations (i.e., 100 repetitions per object and location). The

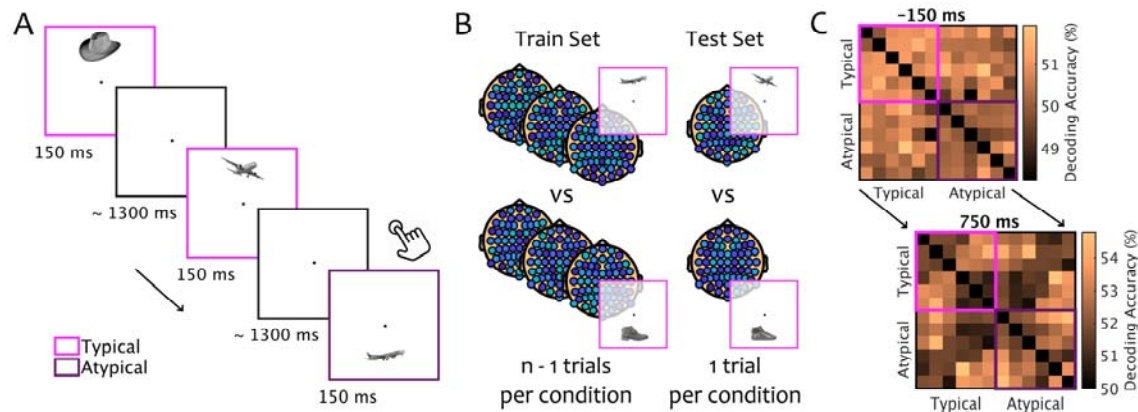207   extended experiment was split into 12 runs.

208

209

210 **Figure 2.** Paradigm and Classification Logic. A) Stimuli were presented for 150 ms in upper

211 or lower visual field locations, corresponding to an object's typical or an atypical location.

212 Participants were instructed to detect occasional one-back repetitions on an object-level

213 (e.g., two airplanes in a row; irrespective of the stimulus location) by pressing a button.

214 Colors indicating the two regularity conditions are shown for illustrative purposes only. B)

215 Multivariate classification was performed on response patterns across all electrodes,

216 separately for each pairwise combination of objects (exemplified here by airplane and

217 shoe in regular locations). The data was split into two sets: a training set consisting of all

218 (but one) trials for each object and a testing set consisting of the two left-out trials. LDA

219 classifiers were repeatedly trained and tested until every trial was left out once and

220 accuracy was averaged across these repetitions. C) The pairwise classification analysis

221 was repeated for each 10 ms time bin, resulting in a 12-by-12 matrix of pairwise

222 classification accuracies (with an empty diagonal) at every time point. To determine

223 differences between typically and atypically positioned objects, pairwise comparisons

224 within the typically placed objects (pink rectangle, upper left) and within the atypically

225 placed objects (purple rectangle, lower right) were averaged and compared (Figure 3).

226

227 *2.4 EEG recording and preprocessing*

228   The EEG was recorded using an EASYCAP 64-channel system and a Brainvision

229   actiCHamp amplifier. The 64 electrodes were arranged in accordance with the standard

230   10-10 system. The data was recorded at a sampling rate of 1000 Hz and filtered online

231   between 0.5 and 70 Hz. For one participant, due to a technical problem, only data from 32

232   electrodes was recorded. All electrodes were referenced online to the Fz electrode.

233   Offline preprocessing was performed in MATLAB, using the FieldTrip toolbox

234   (Oostenveld, Fries, Maris, & Schoffelen, 2011). The continuous EEG data was epoched into

235   trials ranging from 150ms before stimulus onset to 750ms after stimulus onset. Trials

236   containing movement-related artefacts were visually identified and excluded from all

237   analyses. Blink and eye movement artifacts were identified and removed using

238   Independent Components Analysis (ICA) and visual inspection of the resulting

239   components. To increase the signal-to-noise ratio of the classification analyses (Carlson,

240   Tovar, Alink, & Kriegeskorte, 2013), the data was downsampled to 100Hz.

241   *2.5 EEG classification procedure*

242   Multivariate classification analyses were carried out in MATLAB using the

243   CoSMoMVPA toolbox (Oosterhof, Connolly, & Haxby, 2016). Classification was performed

244   separately for each 10ms time bin, resulting in classification time courses with 10 ms

245   resolution. The analysis was performed pairwise, for all possible combinations of the six

246   objects appearing in the two locations. Linear discriminant analysis (LDA) classifiers were

247   always trained and tested on data from two conditions (e.g., an airplane in the upper

248   visual field versus a carpet in the lower visual field), using a leave-one-out partitioning

249   scheme (Figure 2B). The training set consisted of all but one trials for each of the two

250   conditions, while one trial for each of the two conditions was held back and used for

251   classifier testing. This procedure was repeated until every trial was left out once. Classifier

252   performance was averaged across these repetitions. The pairwise decoding analysis

253   resulted in 12-by-12 matrix of decoding accuracies at each time point (reflecting all

254   comparisons between the six objects appearing in the two locations) (Figure 2C).

255   *2.6 Overall classification dynamics*

256   To assess the overall classification dynamics over time, we computed the general

257   discriminability of the twelve different conditions. All pairwise classification accuracies

258   were averaged, revealing a time course of object decoding independently of the

259   positional regularities. This time course of overall classification accuracy was used to

260   define time points of interest at the peaks of the classification time series, where

261   classification performance was particularly pronounced. Using a "region of interest" logic

262   frequently applied in fMRI analyses (Poldrack, 2007), we used these peaks as "time points

263   of interest" to increase the detection power of subsequent analyses.

264   *2.7 Object classification in typical and atypical locations*

265   To determine an effect of positional regularity on object decoding, we compared

266   performance when classifying among typically positioned objects versus among atypically

267   positioned objects. Pairwise classification accuracies were averaged for all comparisons

268   between typically positioned objects (e.g., an airplane in the upper visual field versus a

269   shoe in the lower visual field) and for all comparisons between atypically positioned

270   objects (e.g., a shoe in the upper visual field versus an airplane in the lower visual field)

271   (Figure 2C). Subsequently, the two resulting classification time series (for typically and

272   atypically positioned objects) were compared. To increase the statistical power of this

273   comparison, we specifically focused on the effect of positional regularity at the peaks in

274   overall classification.

275   *2.8 Sensor-space searchlight analysis*

276   To investigate which sensors contributed most to the observed effects, we

277   performed a sensor-space searchlight analysis. For this analysis, the pairwise classification

278   procedure was repeated for neighborhoods of seven adjacent electrodes around each

279   individual electrode; the resulting classification accuracy was then mapped onto a scalp

280   representation. This procedure allowed us to infer the approximate spatial distribution of

281   classification differences between typically and atypically positioned objects. As for one

282   participant only data from 32 electrodes was available, this participant was not included

283   in the searchlight analysis.

284   *2.9 Statistical testing*

285   Statistical testing was done across participants. To identify significant effects

286   across time we used a threshold-free cluster-enhancement procedure (Smith & Nichols,

287   2009) with default parameters. Multiple comparison correction was based on a sign-

288   permutation test (with null distributions created from 10,000 bootstrapping iterations) as

289   implemented in CoSMoMVPA (Oosterhof et al., 2016). The resulting statistical maps were

290   thresholded at $Z > 1.96$ (i.e., $p < .05$). The same procedure was employed for identifying

291   significant sites across electrodes in the sensor-space searchlight analysis. For assessing

292   the significance of effects at the overall classification peaks, repeated-measures ANOVAs

293   and paired *t*-tests were performed. For these tests, we also computed effect size

294   measures: partial η^2 for ANOVAs and Cohen's *d* for *t*-tests. Additionally, for *t*-tests, scaled-

295   information Bayes factors (*BF*) were calculated as an additional measure of evidential

296   value (Rouder, Speckman, Sun, Morey, & Iverson, 2009), where Bayes factors >10 can be

297   considered strong evidence for an effect.

298　**3 Results**

299　*3.1 Temporal dynamics of pairwise object classification*

300　　　In a first step, we characterized the overall response dynamics observed in the

301　pairwise classification analysis, which allowed us to restrict subsequent analyses to time

302　points where classification performance was particularly pronounced. For this, we

303　computed an overall measure of pairwise classification by averaging across all unique off-

304　diagonal elements of the pairwise classification matrices (Figure 2C), resulting in a single

305　classification time series. This analysis revealed robust above-chance classification

306　starting from 70 ms after stimulus onset and prominently peaking at 140 ms and 220 ms

307　(Figure 3A). These two clear peaks in the classification time series were used as time

308　points of interest for subsequent analyses, as we reasoned that differences between

309　typically and atypically positioned objects would be most pronounced at time points at

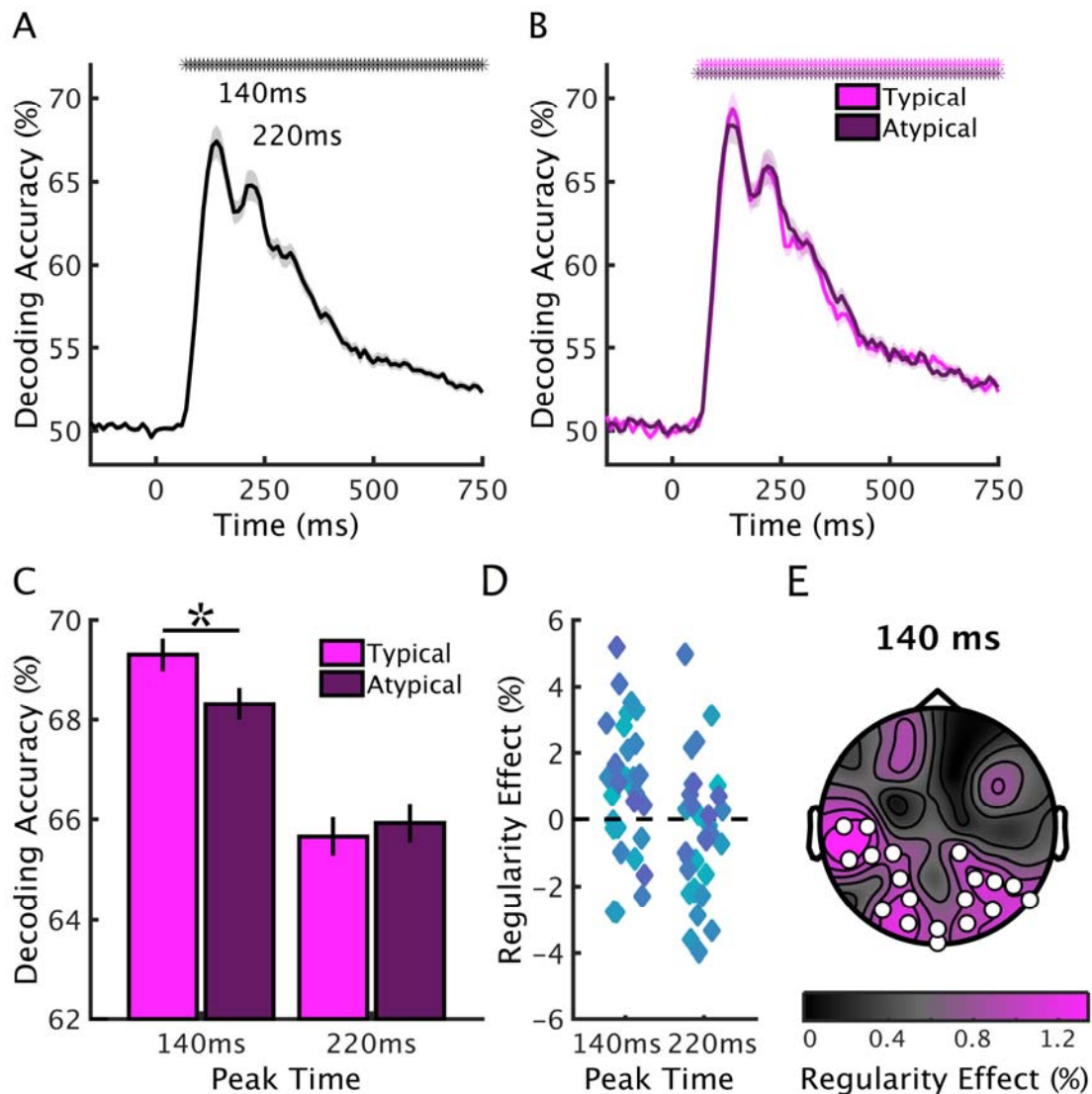310　which objects were most discriminable.

Figure 3. Classification Results. A) Overall pairwise classification performance was computed by averaging all pairwise decoding time series, revealing significant decoding accuracy starting at 70 ms after stimulus onset and peaking at 140 ms and 220 ms. Asterisks above data curves indicate above-chance classification ($p < .05$, corrected for multiple comparisons). The shaded margin reflects standard errors of the mean. B) Classification time series for typically and atypically positioned objects were computed by averaging all pairwise classification time series comparing typically and atypically positioned objects, respectively. Classification of typically and atypically positioned pairs showed comparable temporal dynamics, both peaking at the time points identified in the

321   overall decoding. Asterisks indicate above-chance classification ($p < .05$, corrected for

322   multiple comparisons). The shaded margins reflect standard errors of the mean. C) At the

323   first decoding peak (140 ms), but not the second peak (220 ms), classification was more

324   accurate for typically than for atypically positioned objects. The asterisk indicates a

325   significant difference ($p < .05$). Error bars reflect standard errors of the difference. D)

326   Scatterplots showing the regularity effects across participants, for both peak times. E) A

327   sensor-space classification searchlight revealed that the regularity effect (difference

328   between the classification of typically and atypically positioned objects) at the 140 ms

329   peak is most pronounced in occipital and temporal electrodes. Circles indicate electrodes

330   exhibiting a significant regularity effect ($p < .05$, corrected for multiple comparisons).

331

332   *3.2 Classification of objects when positioned typically and atypically*

333   To test whether neural representations differ for typically and atypically

334   positioned objects, we compared classification performance for all typically and all

335   atypically positioned objects. We averaged all pairwise classification time courses that

336   corresponded to comparisons within regular pairs (e.g., an airplane in the upper visual

337   field versus a boat in the lower visual field) and comparisons within irregular pairs (e.g.,

338   an airplane in the lower visual field versus a boat in the upper visual field) (Figure 3B). The

339   classification time series for typically and atypically positioned objects showed a similar

340   temporal structure and both replicated the peak structure observed in the overall

341   decoding, allowing for a meaningful comparison between typically and atypically

342   positioned objects at the classification peaks. We thus restricted statistical comparisons

343   to two time points of interest: the peak times observed in the overall decoding (140 ms

344   and 220 ms). For the early peak at 140ms, we found higher classification accuracy for

345 typically than for atypically positioned objects, $t(33) = 3.04$, $p = .005$, $d = 0.52$, $BF = 12.06$,

346 while for the later peak at 220ms, no such difference emerged, $t(33) = 0.69$, $p = .495$, $d =$

347 0.12, $BF = 3.37$, interaction with peak time, $F(1,33) = 7.44$, $p = .010$, $\eta^2 = 0.18$ (Figure 3C).

348 This pattern of results suggests that earlier stages of object processing (as reflected in

349 the decoding peak at 140 ms) benefit from typical object locations, while relatively later

350 object representations (at 220 ms after stimulus onset) are not sensitive to positional

351 regularities. This result was replicated in a bootstrapping analysis, where we used

352 independent sub-groups of participants to define peak times and compute the regularity

353 effects (see Supplementary Information, Fig. S1).

354   To estimate the spatial extent of the early regularity effect, we performed a

355 searchlight analysis in sensor space. We repeatedly performed the pairwise classification

356 analysis for neighborhoods of seven adjacent sensors, using only data from the early

357 peak at 140 ms. To quantify the regularity benefit, we then computed the difference

358 between all pairwise comparisons of typically positioned objects and all pairwise

359 comparisons of atypically positioned objects at every sensor location. This analysis

360 revealed a significant regularity effect in posterior and lateral electrodes (19 significant

361 electrode sites) (Figure 3D). This result provides an indication of the tentative cortical

362 source of the regularity effect (within the limits of the restricted spatial resolution of EEG

363 measurements), suggesting that the enhanced classification for regularly positioned

364 objects may originate from visual areas of the occipital and temporal cortices.

365   These results also provide a control for oculomotor influences on our findings. If

366 differential eye-movement patterns were driving the enhanced decoding for typically

367 positioned objects, the effect should be observed in frontal electrodes, too. The absence

368 of a regularity benefit in the frontal electrodes thus provides evidence against eye-

369   movement confounds. Two additional control analyses focused on spatiotemporal

370   response patterns in selected frontal electrodes (see Supplementary Information, Fig.

371   S3); these analyses confirmed the absence of an effect, further rendering eye-movement

372   confounds unlikely.

373   *3.3 Classification within and between locations*

374   Our classification approach collapsed across pairwise comparisons within the

375   same location and between different locations, so that classifiers could rely on

376   information from both an object's identity and its location. Therefore, the classification

377   benefit for typically positioned objects could, in principle, also emerge from a more

378   thorough coding of an object's location. For example, classifiers – in principle – only need

379   to predict the object's location to correctly classify a lamp in the upper visual field versus

380   a shoe in the lower visual field. By contrast, to correctly classify a lamp in the upper visual

381   field versus a hat in the upper visual field, classifiers need to use information about the

382   object. Note that both these comparisons contributed to the classification accuracy of

383   regularly positioned objects in the previous analysis.

384   To test whether a regularity benefit emerges also in situations where classification

385   solely depends on object information, we thus looked at the regularity effect for all

386   comparisons between locations (e.g., an airplane in the upper visual field versus a shoe in

387   the lower visual field) (Figure 4A) and all comparisons within location (e.g., an airplane in

388   the upper visual field versus a hat in the upper visual field) (Figure 4B). If typical

389   positioning enhances location information, a regularity effect should only be found when

390   comparing objects between different locations, but not within a specific location (when

391   location information is eliminated). Finding comparable effects in the within- and

392   between-location comparisons would not support this view, indicating that typical

393   positioning can also enhance processing when only object information is diagnostic.

394   We found a main effect of visual field comparison, $F(1,33) = 365.70$, $p < .001$, $\eta^2 =$

395   0.91, with higher classification accuracies for classifying between locations (where the

396   classifier can use the stimulus' location) than within location (where the classifier has no

397   location information available), and an interaction of the within-between comparison and

398   peak latency, $F(1,33) = 72.30$, $p < .001$, $\eta^2 = 0.69$, with a relatively more pronounced early

399   peak when classifying between locations. Replicating our previous results, the analysis

400   produced a significant peak X regularity interaction, $F(1,33) = 9.83$, $p = .004$, $\eta^2 = 0.23$, with

401   a regularity benefit at the 140 ms peak, $t(33) = 3.15$, $p = .004$, $d = 0.54$, $BF = 15.52$, but not

402   the 220 ms peak, $t(33) = 1.05$, $p = .301$, $d = 0.18$, $BF = 2.50$. Crucially, this pattern of results

403   did not depend on the type of classification (between locations versus within location),

404     $F(1,33) = 2.53$, $p = .121$, $\eta^2 = 0.07$. These results do not provide evidence for a boost of

405   location information, thus suggesting that typical real-world locations enhance visual

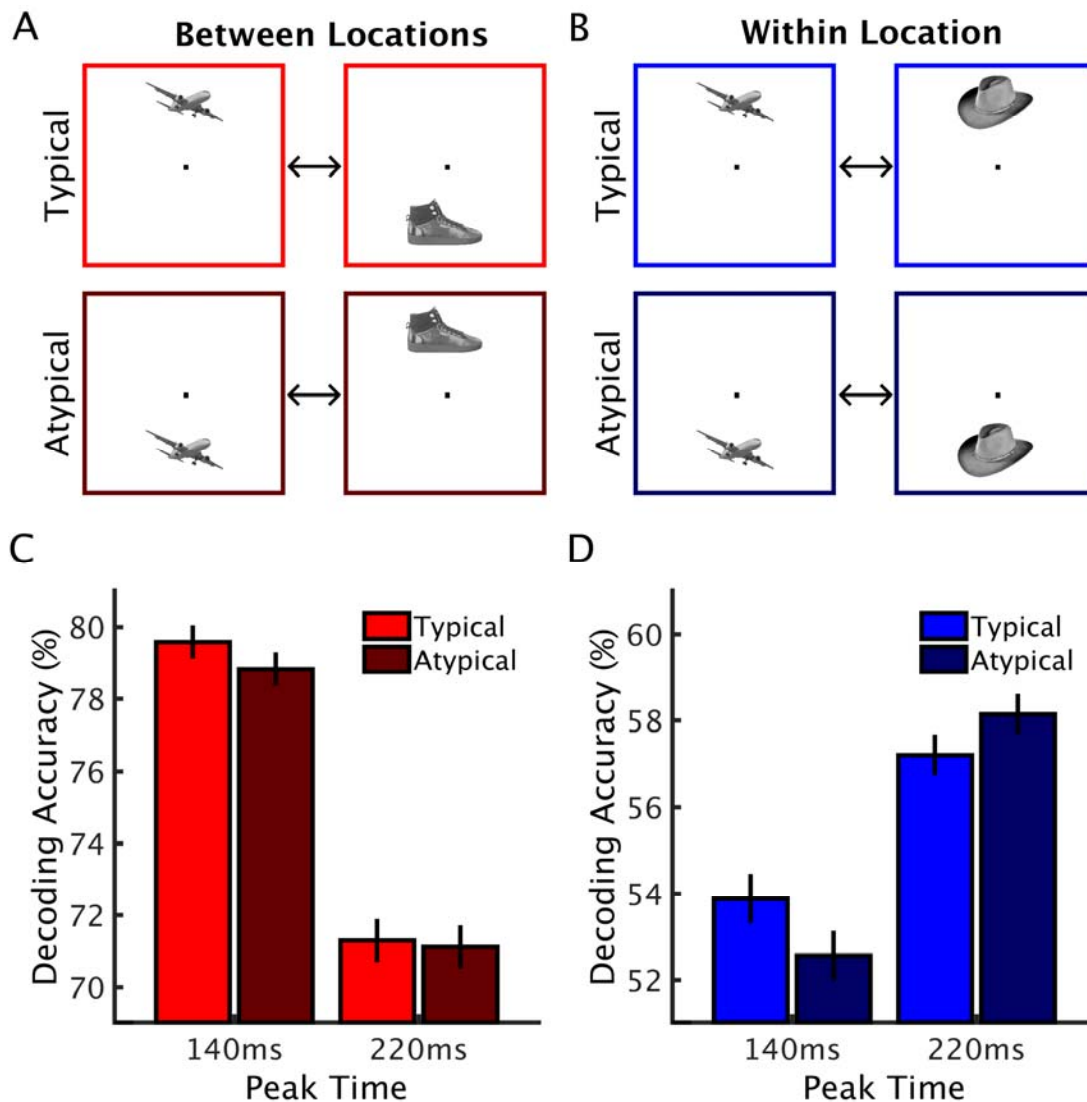406   processing by boosting object identity information.

407

Figure 4. Between-Locations versus Within-Location Classification. We compared peak

decoding for typically and atypically positioned objects separately for comparisons

between different locations (e.g., an airplane in the upper visual field versus a shoe in the

lower visual field) (A) and within the same location (e.g., an airplane in the upper visual

field versus a hat in the upper visual field) (B). For both comparison types, we found a

similar pattern (C, D) with a benefit for typically positioned objects at the 140 ms peak.

Importantly, the patterns of results for between-locations and within-location

classification were statistically indistinguishable (see Results). Error bars reflect standard

errors.

417   ## 4 Discussion

418   *4.1 Summary*

419   Here, we demonstrate that positional regularities contained in real-world scenes

420   impact brain responses to individual objects. Using multivariate classification of EEG data,

421   we show that object coding across the visual field is affected by the typical real-world

422   location of the object. When objects are presented in frequently experienced locations,

423   EEG response patterns at 140 ms after stimulus onset are better discriminable than when

424   the same objects are presented in atypical locations. This early regularity benefit yielded a

425   robust effect size of d = 0.52, and was consistent across subsets of participants (see

426   Supplementary Information). Interestingly, the advantage for typically positioned objects

427   was equally pronounced for classification between locations and within the same

428   location, suggesting that typical positioning boosts object identity information, rather

429   than location information. Using a sensor-space searchlight analysis, we characterize the

430   effect in sensor space, where its most prominent emergence in posterior sensors

431   suggests that typically positioned objects gain an advantage during early perceptual

432   processing.

433   *4.2 Retinotopic priors as a consequence of natural scene structure*

434   How does this retinotopically specific processing benefit emerge from the

435   structure of natural environments? An intuitively appealing explanation lies in the

436   interplay of typical within-scene object locations and gaze patterns during natural

437   viewing. When objects repeatedly occupy specific locations within a scene, they may

438   repeatedly fall into similar parts of the visual field, thereby shaping visual tuning

439   properties. For instance, as lamps mostly appear in the "upper" part of a scene, fixations

440   likely more often land below a lamp than above a lamp. Previous eye-tracking studies can

441 provide some evidence for such a relationship between within-scene locations and

442 retinotopic locations in limited experimental contexts: e.g., when walking along a

443 corridor, lamps most often fall in the upper visual field (Turano et al., 2003). However,

444 gaze patterns in real-world situations are complex and vary as a function of task context

445 (Hayhoe & Ballard, 2005; Land, 2009), making it hard to establish a comprehensive

446 characterization of this relationship. Future large-scale studies that investigate viewing

447 patterns under naturalistic viewing conditions are thus required to empirically establish

448 the connection between typical within-scene locations and typical retinotopic locations

449 with high ecological validity. Without such studies, the association between typical

450 within-scene locations and retinotopic priors – although intuitively compelling – needs to

451 remain somewhat speculative.

452 *4.3 Early stages of object coding are sensitive to typical locations*

453   Our findings demonstrate that visual processing channels are preferentially tuned

454 to specific objects appearing in specific retinotopic locations (Kaiser & Haselhuhn, 2017;

455 Kravitz et al., 2008). Crucially, our EEG classification approach allowed us to pinpoint the

456 latency of this regularity benefit: We demonstrate that object processing at 140 ms after

457 stimulus onset is affected by positional regularity. The timing of the effect suggests that

458 objects appearing in typical visual field locations gain an advantage during early,

459 perceptual processing, rather than through top-down interactions (e.g., via long-range

460 feedback from frontal areas); previous M/EEG studies have suggested that such feedback

461 processes impact visual responses only at later stages, starting shortly before 200 ms

462 (Bar et al., 2006; Fahrenfort, van Leeuwen, Olivers, & Hogendoorn, 2017). As opposed to

463 the difference in early object processing, later representations (at 220 ms after stimulus

464 onset) do not depend on the location of the object. This result concurs with the

465   increasing location tolerance over the time course of object classification, peaking at

466   around 180 ms (Isik et al., 2014), mirroring the increase in receptive field size along the

467   visual stream (Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013).

468   Can the sensitivity to positional regularity be interpreted as a processing

469   advantage for typically positioned objects? Or do the results rather reflect a processing

470   disadvantage for atypically positioned objects? Alternatively, there might be both an

471   experience-based advantage for typically positioned objects, and a disadvantage for

472   atypically positioned objects, caused by a lack of such experience? Our design cannot

473   conclusively disambiguate these possibilities, as it relies on relative effects (i.e.,

474   comparing typical and atypical positioning). In the Supplementary Information, we

475   however provide indirect evidence by analyzing individual-object effects: decoding of

476   typically positioned objects increases when the object has a stronger location prior, while

477   the decoding of atypically positioned objects does not vary as a function of the prior

478   strength (Fig. S2). Although these results provide tentative evidence for a processing

479   advantage for typically positioned objects (rather than a disadvantage for atypically

480   positioned objects), future investigations need to test the different interpretations more

481   explicitly, for example by including objects that are not associated with specific locations.

482   *4.4 Visual versus categorical sources of the regularity benefit*

483   What is the content of the location-specific object representations emerging at

484   140 ms? Peaks in the M/EEG decoding in this time range have been previously associated

485   with visual category processing in object-selective cortex (Cichy, Pantazis, & Oliva, 2014,

486   2016; Carlson et al., 2013). Our searchlight analysis reveals the strongest regularity effect

487   over lateral occipital and temporal electrode sites, suggesting that the effect originates

488   from object-selective visual cortex. Whether processing differences in these object-

489   selective regions reflect genuine category processing differences or whether they reflect

490   differential coding of category-associated visual features is a debated question (Bracci,

491   Ritchie, & op de Beeck, 2017; Peelen & Downing, 2017). While some data suggest that

492   visual properties explain most of the variance in object-selective responses (e.g., Baldassi

493   et al., 2013), a recent MEG decoding study has demonstrated category-selective, rather

494   than visually (shape-) driven, responses from as early as 130 ms after stimulus onset

495   (Kaiser, Azzalini, & Peelen, 2016). To determine if the regularity benefit observed here can

496   be linked to differences in the processing of particular visual features or true categorical

497   processing differences, future studies need to employ stimuli that vary more extensively

498   in their visual characteristics.

499   *4.5 Positional structures beyond person perception*

500   Positional regularities have been studied in humans and non-human primates

501   largely in the context of face and body perception, where parts are arranged in highly

502   predictable configurations (e.g., the features of a human face). fMRI studies in humans

503   have demonstrated that individual face and body parts are processed more efficiently

504   when they appear in typical visual field locations (Chan et al., 2010; de Haas et al., 2016).

505   Single-cell recordings in monkeys demonstrated that location biases can impact cortical

506   responses to face parts as early as 100 ms after stimulus onset (Issa & DiCarlo, 2012),

507   suggesting a benefit at early stages of perceptual processing.

508   Our results complement these findings by showing that such location-specific

509   object processing is not restricted to the face/body domain: The inherent structure of

510   natural scenes can similarly impact early processing (140 ms after stimulus onset) of

511   object information across the visual field. Our findings thus highlight that location-specific

512   tuning in object processing may form a general principle that shapes visual processing

513   mechanisms for spatially predictable information. Future research could test whether

514   regularity structures also affect other domains where the visual input consists of multiple

515   parts that are constrained by spatial regularities. For example, through extensive

516   experience with reading written text, the neural mechanisms for perceiving letters could

517   get tuned to their typical spatial locations within words (Kaiser & Haselhuhn, 2017;

518   Vinckier et al., 2007).

519   *4.6 Positional structures in multiple and individual objects*

520   Natural environments contain positional regularities on different levels, both on

521   the levels of multiple (e.g., a lamps typically hang above tables) and individual objects

522   (e.g., a lamp is typically in the upper visual field). Previous research has primarily focused

523   on the latter: Recent behavioral studies have demonstrated that regularity structures in

524   multi-object arrangement facilitate behavior in capacity-limited visual tasks (Gronau &

525   Shachar, 2014; Kaiser et al., 2014; Kaiser, Stein, & Peelen, 2015; Stein, Kaiser, & Peelen,

526   2015), and neuroimaging studies demonstrated that they enable the brain to integrate

527   information across objects that appear in frequently experienced arrangements (Baeck,

528   Wagemans, & Op de Beeck, 2013; Kaiser & Peelen, 2018; Kaiser et al., 2014).

529   Here, we provide the first evidence that typical regularity structures also impact

530   the neural representation of individual objects. Our finding thus raises the question

531   whether the previously reported regularity effects in multi-object perception can be

532   reduced to the effects of typical individual object location. On a behavioral level, some

533   previous studies oppose this notion by demonstrating that the benefits of multi-object

534   regularities cannot be explained by the relative location of the constituent objects (Kaiser

535   et al., 2014, 2015; Stein et al., 2015). Although these results suggest that positional

536   regularities in multi-object and single-object processing offer complementary benefits,

537   further research is needed. Future studies will need to systematically manipulate

538   positional structures on different levels (from individual objects to multi-object

539   arrangements) to explore how regularities on multiple levels interact on a neural level.

540   *4.7 Object processing in the context of natural scenes*

541   A major challenge for the visual system when processing natural scenes is the

542   large number of individual objects they contain. Surprisingly, however, objects can be

543   rapidly decoded from MEG activity patterns, even when embedded in complex scenes

544   (Brandman & Peelen, 2017; Kaiser, Oosterhof, & Peelen, 2016). Previous

545   electrophysiological studies have investigated how congruent scene context can

546   facilitate cortical processing of an object. Several studies have shown that semantic

547   consistencies (i.e., whether the object is associated with the scene) affect EEG

548   waveforms starting from around 300 ms (Ganis & Kutas, 2003; Mudrik, Lamy, & Deouell,

549   2010; Mudrik, Shalgi, Lamy, & Deouell, 2014; Võ & Wolfe, 2013). More similarly to our

550   study, others have investigated the effects of syntactic consistencies (i.e., whether the

551   object is placed in its typical within-scene location), and found comparably late effects,

552   most prominently between 400 and 600 ms (Demiral, Malcolm, & Henderson, 2012; Võ &

553   Wolfe, 2013). Such effects of syntactic consistencies have been linked to efficient

554   behavioral performance in complex multi-object scenes (Draschkow & Võ, 2017).

555   How can the early effect observed here be reconciled with the late effects found

556   for typical object placement within a scene? Studies on object-scene consistencies differ

557   in two important ways from our study. First, in these studies participants are cued to look

558   at the object's location within the scene, whereby retinotopic object location is

559   purposefully eliminated. Second, these studies measure overall waveform changes

560   between congruent and incongruent conditions, which may reflect more general, post-

561   perceptual signatures of consistency processing or object-scene interaction. By contrast,

562   our analysis used highly sensitive decoding methods to explicitly focus on object

563   discriminability within typically and atypically positioned objects, so that our results are

564   best understood as a rapid signature of experience-based tuning of the visual

565   architecture. Our findings and the findings of these previous studies thus highlight

566   different facets of the processing of regularly structured environments, where both the

567   early effects of typical retinotopic locations and the later effects of typical within-scene

568   positioning may contribute to efficient real-world perception.

569   *4.8 Conclusion*

570   Together, our results demonstrate a general principle of object coding in human

571   visual cortex: Information about an object's location and its identity are processed

572   interactively, where objects appearing in their typical retinotopic locations are more

573   efficiently coded than objects appearing in atypical retinotopic locations. This processing

574   advantage manifests in object decoding 140 ms after stimulus onset, suggesting that

575   early object representations are tied to the object's extensively experienced visual-field

576   location. This finding can provide a novel explanation for the efficient parsing of complex

577   real-world environments, which contain large numbers of typically positioned objects.

578

579   **Acknowledgements**

581 **References**

582 Baeck, A., Wagemans, J., & Op de Beeck, H. P. (2013). The distributed representation of

583 random and meaningful object pairs in human occipitotemporal cortex: the

584 weighted average as a general rule. *Neuroimage, 70*, 37-47.

585 Baldassi, C., Alemi-Neissi, A., Pagan, M., DiCarlo, J. J., Zecchina, R., & Zoccolan, D. (2013).

586 Shape similarity, better than semantic membership, accounts for the structure of

587 visual object representations in a population of monkey inferotemporal neurons.

588 *PLoS Computational Biology, 9*, e1003167.

589 Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*, 617-629.

590 Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen,

591 M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down

592 facilitation of visual recognition. *Proceedings of the National Academy of Sciences,*

593 *U.S.A., 103*, 449-454.

594 Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: detecting

595 and judging objects undergoing relational violations. *Cognitive Psychology, 14*, 143-

596 177.

597 Bracci, S., Ritchie, J. B., & op de Beeck, H. P. (2017). On the partnership between neural

598 representations of object categories and visual features in the ventral visual

599 pathway. *Neuropsychologia, 105*, 153-164.

600 Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

601 Brandman, T., & Peelen, M. V. (2017). Interaction between scene and object processing

602 revealed by human fMRI and MEG decoding. *Journal of Neuroscience, 37*, 7700-

603 7710.

604   Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal

605   resolution decoding of object position and category. *Journal of Vision, 11, 9.*

606   Carlson, T. A., Tovar D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics

607   of object vision: the first 1000 ms. *Journal of Vision, 13, 1-19.*

608   Chan, A. W., Kravitz, D. J., Truong, S., Arizpe, J., & Baker, C.I. (2010). Cortical

609   representations of bodies and faces are strongest in commonly experienced

610   configurations. *Nature Neuroscience, 13, 417-418.*

611   Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4,*

612   170-178.

613   Cichy, R. M., Chen, Y., & Haynes, J. D. (2011). Encoding the identity and location of objects

614   in human LOC. *Neuroimage, 54, 2297-2307.*

615   Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space

616   and time. *Nature Neuroscience, 17, 455-462.*

617   Cichy, R. M., Pantazis, D., & Oliva, A. (2016). Similarity-based fusion of MEG and fMRI

618   reveals spatio-temporal dynamics in human cortex during visual object

619   recognition. *Cerebral Cortex, 26, 3563-3579.*

620   Contini, E. W., Wardle, S. G., & Carlson, T. A. (2017). Decoding the time-course of object

621   recognition in the human brain: From visual features to categorical decisions.

622   *Neuropsychologia, 105, 165-176.*

623   de Haas, B., Schwarzkopf, D. S., Alvarez, I., Lawson, R. P., Henriksson, L., Kriegeskorte, N.,

624   & Rees, G. (2016). Perception and processing of faces in the human brain is tuned

625   to typical facial feature locations. *Journal of Neuroscience, 36, 9289-9302.*

626   Demiral, S. B., Malcolm, G. L., & Henderson, J. M. (2012). ERP correlates of spatially

627   incongruent object identification during scene viewing: contextual expectancy

628   versus simultaneous processing. *Neuropsychologia, 50*, 1271-1285.

629   Draschkow, D., & Võ, M. L.-H. (2017). Scene grammar shapes the way we interact with

630   objects, strengthens memories, and speeds search. *Scientific Reports, 7*, 16471.

631   Fahrenfort, J. J., van Leeuwen, J., Olivers, C. N. L., & Hogendoorn, H. (2017). Perceptual

632   integration without conscious access. *Proceedings of the National Academy of*

633   *Sciences, U.S.A., 114*, 3744-3749.

634   Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object

635   identification. *Cognitive Brain Research, 16*, 123-144.

636   Golomb, J. D., & Kanwisher, N. (2012). Higher level visual cortex represents retinotopic,

637   not spatiotopic, object location. *Cerebral Cortex, 22*, 2794-2810.

638   Gronau, N., & Shachar, M. (2014). Contextual integration of visual objects necessitates

639   attention. *Attention, Perception & Psychophysics, 76*, 695-714.

640   Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an

641   organizing principle for human high-order object areas. Neuron, *34*, 479-490.

642   Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive*

643   *Sciences, 9*, 188-194.

644   Hemond, C. C., Kanwisher, N., & Op de Beeck, H. P. (2007). A preference for contralateral

645   stimuli in human object- and face-selective cortex. *PLoS One, 2*, e574.

646   Henriksson, L., Mur, M., & Kriegeskorte, N. (2015). Faciotopy – A face-feature like

647   topology in the human occipital face area. *Cortex, 72*, 156-167.

648   Hong, H, Yamins, D. L. K., Majaj, N. J., & DiCarlo, J. J. (2016). Explicit information for

649   category-orthogonal object properties increases along the ventral visual stream.

650   *Nature Neuroscience, 19*, 613-622.

651   Isik, L., Meyers, E. M., Leibo, J. Z., & Poggio, T. (2014). The dynamics of invariant object

652   recognition in the human visual system. *Journal of Neurophysiology, 111*, 91-102.

653   Issa, E. B., & DiCarlo, J. J. (2012). Precedence of the eye region in neural processing of

654   faces. *Journal of Neuroscience, 32*, 16666-16682.

655   Kaiser, D., Azzalini, D. C., & Peelen, M. V. (2016). Shape-independent object category

656   responses revealed by MEG and fMRI decoding. *Journal of Neurophysiology, 115*,

657   2246-2250.

658   Kaiser, D., & Haselhuhn, T. (2017). Facing a regular world: How spatial object structure

659   shapes visual processing. *Journal of Neuroscience, 37*, 1965-1967.

660   Kaiser, D., Oosterhof, N. N., & Peelen, M. V. (2016). The neural dynamics of attentional

661   selection in natural scenes. *Journal of Neuroscience, 36*, 10522-10528.

662   Kaiser. D., & Peelen, M. V. (2018). Transformation from independent to integrative coding

663   of multi-object arrangements in human visual cortex. *Neuroimage, 169*, 334-341.

664   Kaiser, D., Stein, T., & Peelen, M. V. (2014). Object grouping based on real-world

665   regularities facilitates perception by reducing competitive interactions in visual

666   cortex. *Proceedings of the National Academy of Sciences, U.S.A., 111*, 11217-11222.

667   Kaiser, D., Stein, T., & Peelen, M. V. (2015). Real-world spatial regularities affect visual

668   working memory for objects. *Psychonomic Bulletin & Review, 22*, 1784-1790.

669   Kravitz, D. J., Kriegeskorte, N., & Baker, C. I. (2010). High-level visual object

670   representations are constrained by position. *Cerebral Cortex, 20*, 2916-2925.

671   Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The

672   ventral visual pathway: an expanded neural framework for the processing of

673   object quality. *Trends in Cognitive Sciences, 17,* 26-49.

674   Kravitz, D. J., Vinson, L. D., & Baker, C. I. (2008). How position dependent is visual object

675   recognition? *Trends in Cognitive Sciences, 12,* 114-122.

676   Land, M. F. (2009). Vision, eye movements, and natural behavior. *Visual Neuroscience, 26,*

677   51-62.

678   Mudrik, L., Lamy, D., & Deouell, L. Y. (2010). ERP evidence for context congruity effects

679   during simultaneous object-scene processing. *Neuropsychologia, 48,* 507-517.

680   Mudrik, L., Shalgi, L., Lamy, D., & Deouell, L. Y. (2010). ERP correlates of spatially

681   incongruent object identification during scene viewing: contextual expectancy

682   versus simultaneous processing. *Neuropsychologia, 56,* 447-458.

683   Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in*

684   *Cognitive Sciences, 11,* 520-527.

685   Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). Fieldtrip: open source

686   software for advances analysis of MEG, EEG, and invasive electrophysiological

687   data. *Computational Intelligence and Neuroscience, 2011,* 156869.

688   Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: multi-modal

689   multivariate pattern analysis of neuroimaging data in Matlab / GNU Octave.

690   *Frontiers in Neuroinformatics, 10,* 20.

691   Peelen, M. V., & Downing, P. E. (2017) Category selectivity in human visual cortex: Beyond

692   visual object recognition. *Neuropsychologia, 105,* 177-183.

693   Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social Cognitive and Affective*

694   *Neuroscience, 2,* 67-70.

695   Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t-tests

696   for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16,*

697   225-237.

698   Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: a database

699   and web-based tool for image annotation. *International Journal of Computer Vision,*

700   *77,* 157-173.

701   Schwarzlose, R. F., Swisher, J. D., Dang, S., & Kanwisher, N. (2008). The distribution of

702   category and location information across object-selective regions in human visual

703   cortex. *Proceedings of the National Academy of Sciences, U.S.A., 105,* 4447-4452.

704   Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing

705   problems of smoothing, threshold dependence and localization in cluster

706   interference. *Neuroimage, 66,* 215-222.

707   Stein, T., Kaiser, D., & Peelen, M. V. (2015). Interobject grouping facilitates visual

708   awareness. *Journal of Vision, 15,* 10.

709   Turano, K. A., Geruschat, D. R., & Baker, F. H. (2003). Oculomotor strategies for the

710   direction of gaze tested with a real-world activity. *Vision Research, 43,* 333-346.

711   Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007).

712   Hierarchical coding of letter strings in the ventral stream: dissecting the inner

713   organization of the visual word-form system. *Neuron, 55,* 143-156.

714   Võ, M. L.-H., & Wolfe, J. M. (2013). Differential ERP signatures elicited by semantic and

715   syntactic processing in scenes. *Psychological Science, 24,* 1816-1823.

716   Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010).

717   Controlling low-level image properties: The SHINE toolbox. *Behavior Research*

718   *Methods, 42,* 671-684.

719   Wolfe, J. M., Võ, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes

720   involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15*, 77-

721   84.