

1 Now you see it, now you don't: Overlapping neural representations for the position of visible and  
2 invisible objects

3 Amanda K. Robinson<sup>a,b,c</sup>, Tijl Grootswagers<sup>a,b,c</sup>, Sophia M. Shatek<sup>a</sup>, Jack Gerboni<sup>a</sup>, Alex O. Holcombe<sup>a</sup>,  
4 Thomas A. Carlson<sup>a,b</sup>

5

6 <sup>a</sup> School of Psychology, University of Sydney, NSW, 2006, Australia

7 <sup>b</sup> Perception in Action Research Centre, Macquarie University, Sydney, NSW, Australia

8 <sup>c</sup> Department of Cognitive Science, Macquarie University, NSW, 2109, Australia

9

10 **Abstract**

11 Humans can covertly track the position of an object, even if the object is temporarily occluded. What are  
12 the neural mechanisms underlying our capacity to track moving objects when there is no physical stimulus  
13 for the brain to track? One possibility is that the brain “fills-in” information about invisible objects using  
14 internally generated representations similar to those generated by feed-forward perceptual mechanisms.  
15 Alternatively, the brain might deploy a higher order mechanism, for example using an object tracking  
16 model that integrates visual signals and motion dynamics (Kwon et al., 2015). In the present study, we  
17 used electroencephalography (EEG) and time-resolved multivariate pattern analyses to investigate the  
18 spatial processing of visible and invisible objects. Participants tracked an object that moved in discrete  
19 steps around fixation, occupying six consecutive locations. They were asked to imagine that the object  
20 continued on the same trajectory after it disappeared and move their attention to the corresponding  
21 positions. Time-resolved decoding of EEG data revealed that the location of the visible stimuli could be  
22 decoded shortly after image onset, consistent with early retinotopic visual processes. For processing of  
23 unseen/invisible positions, the patterns of neural activity resembled stimulus-driven mid-level visual  
24 processes, but were detected earlier than perceptual mechanisms, implicating an anticipatory and more  
25 variable tracking mechanism. Monitoring the position of invisible objects thus utilises similar perceptual

26 processes as processing objects that are actually present, but with different temporal dynamics. These  
27 results indicate that internally generated representations rely on top-down processes, and their timing is  
28 influenced by the predictability of the stimulus. All data and analysis code for this study are available at  
29 <https://osf.io/8v47t/>.

30 **Introduction**

31 Internally-generated representations of the world, as opposed to stimulus-driven feedforward  
32 representations, are important for day-to-day tasks such as constructing a mental map to give a stranger  
33 directions, remembering where you last saw a lost item, or tracking the location of a car that becomes  
34 occluded by another vehicle. In these cases, there is little or no relevant perceptual input, yet the brain  
35 successfully constructs a picture of relevant visual features such as object form and spatial position. Such  
36 internally-generated representations have been studied with tasks involving imagery, mental rotation,  
37 and perception of occluded objects. It is clear that internally-generated representations rely on similar  
38 brain regions to stimulus-driven perceptual representations (Lee et al., 2012; Reddy et al., 2010) but they  
39 appear to have different temporal dynamics (Dijkstra et al., 2018), raising the question of how exactly  
40 these internal representations are formed.

41

42 Top-down processing appears to play an important role in generating internally representations. Current  
43 theories of mental imagery are based on similarities between perception and imagery, with a greater  
44 focus on bottom-up processing in perception and top-down processing in imagery (for review, see  
45 Pearson, 2019). Neuroimaging work has shown increases in brain activation within early visual cortical  
46 regions when participants engage in imagery, in a similar way to viewing the same stimuli (Kosslyn et al.,  
47 1993; Le Bihan et al., 1993), but there is more perception-imagery overlap in higher level brain regions  
48 such as ventral temporal cortex (Lee et al., 2012; Reddy et al., 2010). Imagery involves greater flow of  
49 information from fronto-parietal to occipital regions than perception, indicating that top-down or  
50 feedback-like processes mediate internally generated representations (Dentico et al., 2014; Dijkstra et al.,  
51 2017; Mechelli, 2004). Consistent with this account, recent work using magnetoencephalography and  
52 time-resolved decoding showed that imagery of faces and houses involves similar patterns of activation  
53 as viewing those stimuli, but with different temporal dynamics (Dijkstra et al., 2018). In the Dijkstra et al.  
54 (2018) study, imagery-related processing was delayed and more diffuse than perception, which showed

55 multiple distinct processing stages. Together, these results suggest that imagery originates in higher-level  
56 brain regions rather than involving feed-forward visual processes from V1.

57

58 One aspect that is likely to affect the top-down generation of internal representations is the ability to  
59 predict aspects of the stimulus in advance, for example when objects become occluded. The processes  
60 underlying the representation of occluded objects may be closely related to those in conventional imagery  
61 tasks (Nanay, 2010). However, there are some important differences between imagery and occlusion.  
62 Imagery can be prompted from either long term memory or working memory, which involve different  
63 patterns of brain activation (Ishai, 2002), whereas representations in conditions of occlusion often have  
64 some sensory support, such as from a fragment of the object not occluded or full view of the object  
65 immediately before occlusion. One possibility is that internally generated representations utilise the same  
66 brain networks as perceptual representations but the temporal dynamics vary with the ability to predict  
67 and anticipate details of the stimulus to be generated.

68

69 Tracking the position of a predictably moving object is a common task that may share some top-down  
70 processes with static imagery tasks. In particular, prediction is likely to play an important role in both  
71 imagery and visual tracking. The ability to predict the movement of a stimulus influences perceptual  
72 processing during visual tracking; Hogendoorn & Burkitt (2018) measured EEG from participants who  
73 viewed an apparent motion stimulus that was predictable or unpredictable in its motion trajectory.  
74 Position-specific representations 80-90ms after stimulus onset were unaffected by the predictability of  
75 the motion, but a later stage of processing (typically 140-150ms after a stimulus is presented) was pre-  
76 activated for predictable relative to random sequences by approximately 16ms (Hogendoorn & Burkitt,  
77 2018). Predictability therefore has a marked effect on the temporal dynamics of spatial representations  
78 for visible stimuli. For an object appearing in an unpredictable location, the resulting position  
79 representation must be a combination of the internal representation of the expected location and the

80 stimulus-driven response to the actual object location. Disentangling how expected stimulus position is  
81 represented in the brain, the *internal* spatial representation, from a stimulus-driven response, is an  
82 important next step in understanding how and when internal representations are formed. Anticipatory  
83 mechanisms are likely to influence internally generated spatial representations, but might interact with  
84 other effects, for example the delayed processes observed during imagery (Dijkstra et al., 2018).

85

86 In the current study, to understand the nature of internal representations in the brain, we investigated  
87 the neural processes underlying visual tracking for visible and invisible objects. Participants covertly  
88 tracked the position of a simple moving stimulus and kept tracking its imaginary trajectory after it  
89 disappeared. Using invisible objects allowed us to assess the temporal dynamics of internal  
90 representations during object tracking in the absence of a stimulus-driven response. EEG and time-  
91 resolved multivariate pattern analysis were used to assess the position-specific information contained  
92 within the neural signal during visible and invisible stimulus presentations. We successfully decoded the  
93 position of the stimuli from all phases of the task. Our results show that the visible and invisible stimuli  
94 evoked the same neural response patterns, but with very different temporal dynamics. These findings  
95 suggest that overlapping mid- and high-level visual processes underlie perceptual and internally  
96 generated representations of spatial location, and that these are pre-activated in anticipation of a  
97 stimulus.

98

## 99 Methods

100 All stimuli, data and analysis code are available at <https://osf.io/8v47t/>. The experiment consisted of two  
101 types of sequences: a template pattern estimator and the experimental task. The pattern estimator used  
102 unpredictable stimulus sequences to obtain position-specific EEG signals that were unlikely to be affected  
103 by eye-movements. These were subsequently used to detect position signals in the experimental task.

104

105 ***Participants***

106 Participants were 20 adults recruited from the University of Sydney (12 females; age range 18-52 years)  
107 in return for payment or course credit. The study was approved by the University of Sydney ethics  
108 committee and informed consent was obtained from all participants. Four participants were excluded  
109 from analyses due to excessive eye movements during the template pattern estimator sequences.

110

111 ***Stimuli and design***

112 While participants maintained fixation in the centre of the monitor, a stimulus appeared in six distinct  
113 positions 4 degrees of visual angle from fixation. The stimulus was a black circle with a diameter of 3  
114 degrees of visual angle. Six unfilled circles acted as placeholders, marking all possible positions throughout  
115 the trial. Every stimulus presentation was accompanied by a 1000 Hz pure tone presented for 100 ms via  
116 headphones. All stimuli were presented using Psychtoolbox (Brainard, 1997; Kleiner et al., 2007; Pelli,  
117 1997) in MATLAB. In total, there were 8 blocks of trials, each of which contained two template pattern  
118 estimator sequences and 36 experimental task sequences.

119

120 ***Template pattern estimator***

121 The template pattern estimator sequences were designed to extract stimulus-driven position-specific  
122 neural patterns from the EEG signal. Participants viewed 16 pattern estimator sequences (2 per block),  
123 each of which consisted of 10 repetitions of the 6 stimulus positions (Figure 1a). The order of stimuli was  
124 randomised to ensure that for a given stimulus position, the preceding and following stimuli would not  
125 be predictive of that position; for example, comparing the neural patterns evoked by positions 1 and 2  
126 could not be contaminated by preceding and following stimuli because they could both be preceded and  
127 followed by all six positions. Each stimulus was shown for 100ms and was followed by an inter-stimulus  
128 interval of 200ms. Onset of the stimulus was accompanied by a 100ms tone. Participants were instructed  
129 to passively view the stimuli without moving their eyes from the fixation cross in the centre of the screen.

130

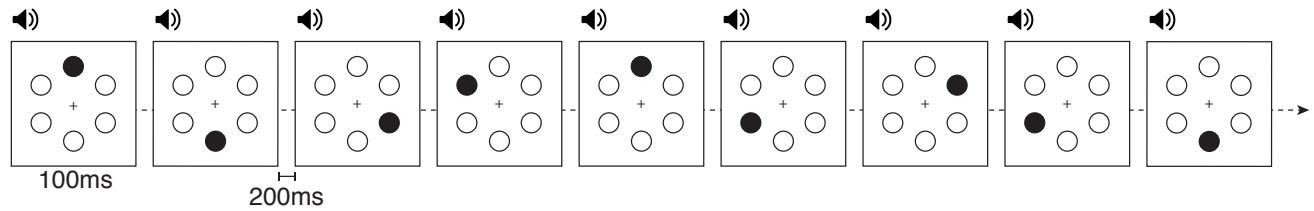
131 The stimuli were presented in unpredictable patterns so there was no regularity in the positions of the  
132 previous or following stimuli to contribute to the neural patterns extracted for each position. Additionally,  
133 the random sequences ensured that any eye movements would be irregular and thus unlikely to  
134 contribute to the extracted neural signal. Previous work has shown that even the fastest saccades typically  
135 take at least 100ms to initiate (Fischer & Ramsperger, 1984). Furthermore, eye movements do not appear  
136 to affect decoding of magnetoencephalography data until 200ms after a lateralised stimulus is presented  
137 (Quax et al., 2019). Our 100ms stimulus duration was therefore unlikely to generate consistent eye  
138 movements that would affect the early, retinotopic EEG signal of stimulus position.

139

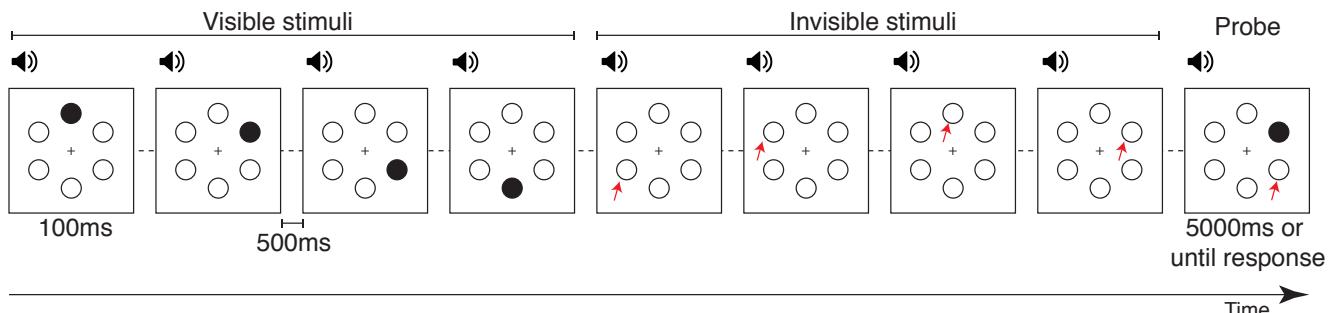
140 To assess whether participants complied with the fixation instruction, we assessed the EEG signal from  
141 electrodes AF7 and AF8 (located near the left and right eye, respectively) as a proxy for electrooculogram  
142 measurements. We calculated the standard deviation of the AF7 and AF8 signals across each of the 16  
143 sequences and then averaged the deviation for the two electrodes. If a participant's average median  
144 deviation across the 16 sequences exceeded 50 $\mu$ V, that individual was considered to be moving their eyes  
145 or blinking too often, resulting in poor signal. An amplitude threshold of 100  $\mu$ V is commonly used to  
146 designate gross artefacts in EEG signal (Luck, 2005), so we adopted an arbitrary standard deviation  
147 threshold of 50  $\mu$ V (50% of the typical amplitude threshold) to indicate that there were too many artefacts  
148 across the entire pattern estimator sequences. Four participants exceeded this standard deviation  
149 threshold ( $M = 72.72\mu$ V, range = 63.93-82.70 $\mu$ V) and were excluded from all analyses. For each of the  
150 remaining 16 participants, the median deviation was well below this threshold ( $M = 25.92\mu$ V,  $SD = 5.64\mu$ V,  
151 range = 16.06-37.62 $\mu$ V). Thus, the four excluded participants had far more signal artefacts (probably  
152 arising from eye movements) than the other participants.

153

A. Template pattern estimator



B. Tracking task



154

155 Figure 1. Stimuli and design. a) Template pattern estimator. Participants passively viewed rapid sequences in which a black  
156 circle stimulus appeared in six locations in random order. A tone accompanied every stimulus onset. b) Tracking task. The  
157 stimulus was presented in different locations in predictable sequences. After 4-6 visible locations, participants had to track the  
158 location of the “invisible” stimulus by imagining the continuation of the sequence. A tone accompanied every stimulus onset.  
159 During the 4-6 “imagined” positions, the auditory stimulus continued at the same rate, but only the six placeholder locations  
160 were shown. At the end of the sequence, a probe appeared, and participants had to respond if it was in the expected position  
161 or whether it was trailing or leading the sequence. This example shows a clockwise sequence with trailing probe. Red arrows  
162 (not shown in experiment) designate the expected position of the invisible stimulus.

163

164 *Tracking task*

165 For the experimental task, participants viewed sequences consisting of 4-6 visible stimuli and 4-6  
166 “invisible” presentations simulating occluded stimuli (Figure 1b). The positions of the visible stimuli were  
167 predictable, presented in clockwise or counter-clockwise sequences. Participants were asked to covertly  
168 track the position of the stimulus, and to continue imagining the sequence of positions when the stimulus  
169 was no longer visible. At the end of each sequence, there was a 1000 ms blank screen followed by a probe  
170 stimulus that was presented in one of the 6 locations. Participants categorised this probe as either (1)  
171 *trailing*: one position behind in the sequence, (2) *expected*: the correct location, or (3) *leading*: one

172 position ahead in the sequence. Participants responded using the Z, X or C keys on a keyboard,  
173 respectively. Each response was equally likely to be correct, so chance performance was 33.33%.

174

175 **EEG recordings and preprocessing**

176 EEG data were continuously recorded from 64 electrodes arranged in the international 10–10 system for  
177 electrode placement (Oostenveld & Praamstra, 2001) using a BrainVision ActiChamp system, digitized at  
178 a 1000-Hz sample rate. Scalp electrodes were referenced to Cz during recording. EEGLAB (Delorme &  
179 Makeig, 2004) was used to pre-process the data offline, where data were filtered using a Hamming  
180 windowed sinc FIR filter with highpass of 0.1Hz and lowpass of 100Hz and then downsampled to 250Hz  
181 as in our previous work (Grootswagers et al., 2019; Robinson et al., 2019). Epochs were created for each  
182 stimulus presentation ranging from -200 to 1000ms relative to stimulus onset. No further preprocessing  
183 steps were applied.

184

185 **Decoding analyses**

186 An MVPA decoding pipeline (Grootswagers et al., 2017) was applied to the EEG epochs to investigate  
187 position representations of visible and invisible stimuli. All steps in the decoding analysis were  
188 implemented in CoSMoMVPA (Oosterhof et al., 2016). A leave-one-block-out (i.e., 8-fold) cross-validation  
189 procedure was used for all time-resolved analyses. A linear discriminant analysis classifier was trained  
190 using the template pattern estimator data to distinguish between all pairs of positions. The classifier was  
191 trained with balanced numbers of trials per stimulus position from the template pattern estimator  
192 sequences. The classifier was then tested separately on the visible and invisible positions in the  
193 experimental task. This provided decoding accuracy over time for each condition. At each time point,  
194 mean pairwise accuracy was tested against chance (50%). Importantly, because all analyses used the  
195 randomly-ordered template pattern estimator data for training the classifier, above chance classification  
196 was very unlikely to arise from the predictable sequences or eye movements in the experimental task. For

197 the tracking task, all sequences were included in the decoding analyses regardless of whether the  
198 participant correctly classified the position of the probe (i.e., correct and incorrect sequences were  
199 analysed). When only correct trials were included, the trends in the results remained the same (see Figure  
200 S1, <https://osf.io/8v47t/>).

201

202 To assess whether neighbouring stimulus positions evoked more similar neural responses, we also  
203 calculated decoding accuracy as a function of the distance between position pairs. Each position pair had  
204 a radial distance of 60°, 120° or 180° apart. There were six pairs with a distance of 60° (e.g., position 1 vs  
205 position 2, position 2 vs position 3), six pairs with a distance of 120° (e.g., position 1 vs position 3, position  
206 2 vs position 4), and three pairs with a distance of 180° (directly opposing each other, e.g., position 1 vs  
207 position 4, position 2 vs position 5). Decoding accuracy for each pair distance was calculated as the mean  
208 of all relevant pair decoding and compared to chance (50%).

209

210 As a final set of analyses, time generalisation (King & Dehaene, 2014) was used to assess whether the  
211 patterns of informative neural activity occurred at the same times for the pattern localiser and the visible  
212 and invisible stimuli on the tracking task. Classification was performed on all combinations of time points  
213 from the pattern estimator epochs and the visible or invisible epochs. The classifier was trained on all  
214 trials from the localiser sequences and tested on visible and invisible stimulus positions. To reduce  
215 computation time, instead of the 15 pairwise tests conducted for the time-resolved decoding analyses,  
216 we performed six-way position decoding for the time generalization analyses, so chance was 16.66%.

217

## 218 **Statistical inference**

219 To assess the evidence that decoding performance differed from chance, we calculated Bayes factors  
220 (Dienes, 2011; Jeffreys, 1961; Kass & Raftery, 1995; Rouder et al., 2009; Wagenmakers, 2007). A JZS prior  
221 (Rouder et al., 2009) was used with a scale factor of 0.707 to test the alternative hypothesis of above-

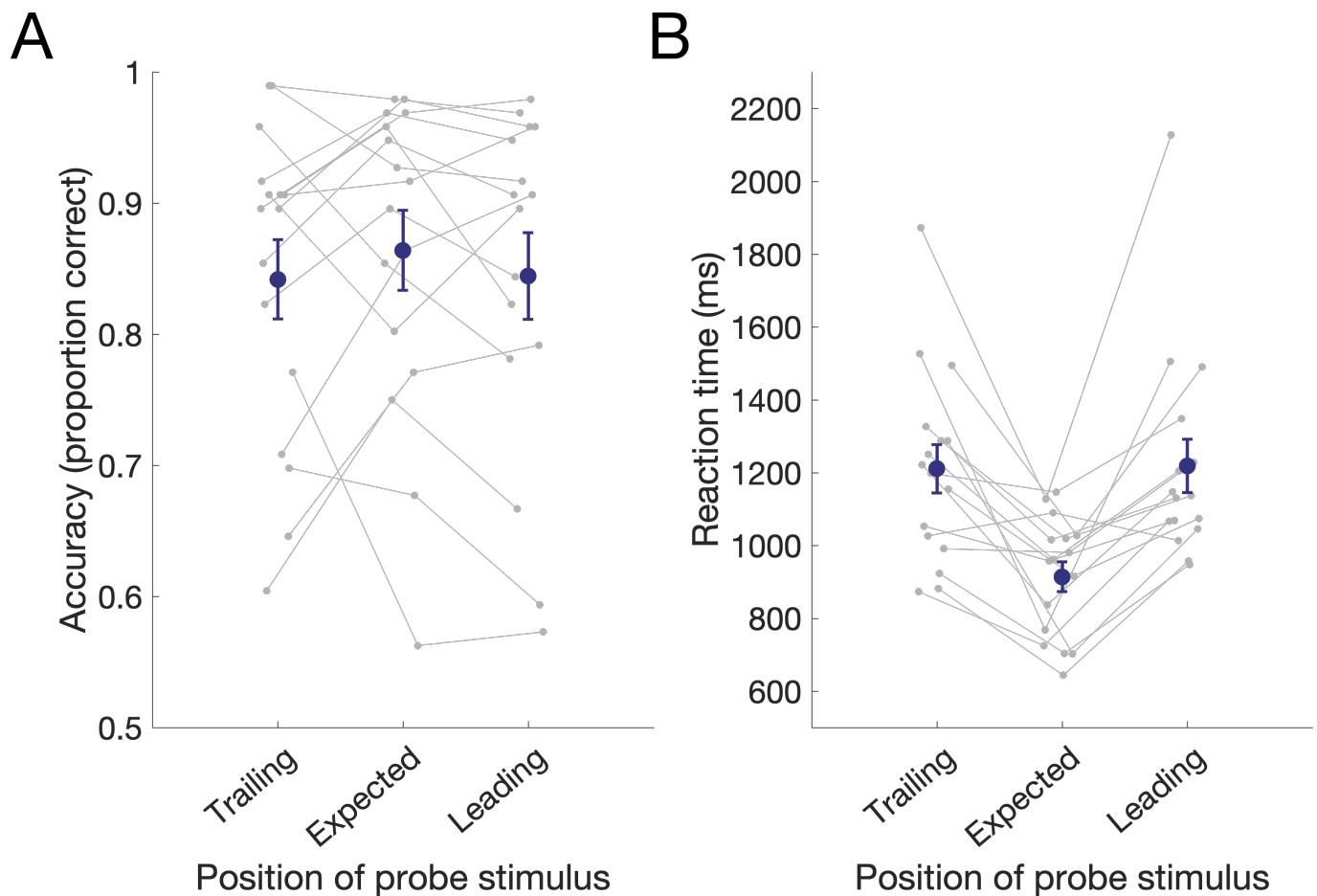
222 chance decoding (Jeffreys, 1961; Rouder et al., 2009; Wetzels & Wagenmakers, 2012; Zellner & Siow,  
223 1980). The Bayes factor (BF) indicates the probability of obtaining the group data given the alternative  
224 hypothesis relative to the probability of the data assuming the null hypothesis is true. We used a threshold  
225 of  $BF > 3$  as evidence for the alternative hypothesis, and  $BF < 1/3$  as evidence in favour of the null  
226 hypothesis (Jeffreys, 1961; Kass & Raftery, 1995; Wetzels et al., 2011). BFs that lie between those values  
227 indicate insufficient evidence to favour a hypothesis.

228

229 **Results**

230 *Behavioural results*

231 Participants performed well on the tracking task. Mean accuracy was high for all probe positions (Fig 2a),  
232 and response time was faster for the expected probe position relative to the unexpected probe positions  
233 (trailing or leading) (Fig 2b). These results indicate that on most trials participants knew where the probe  
234 was meant to appear, which required tracking the expected location of the object. Therefore, participants  
235 allocated their attention appropriately to the expected position of the stimulus during the invisible  
236 portion of the tracking task.



237

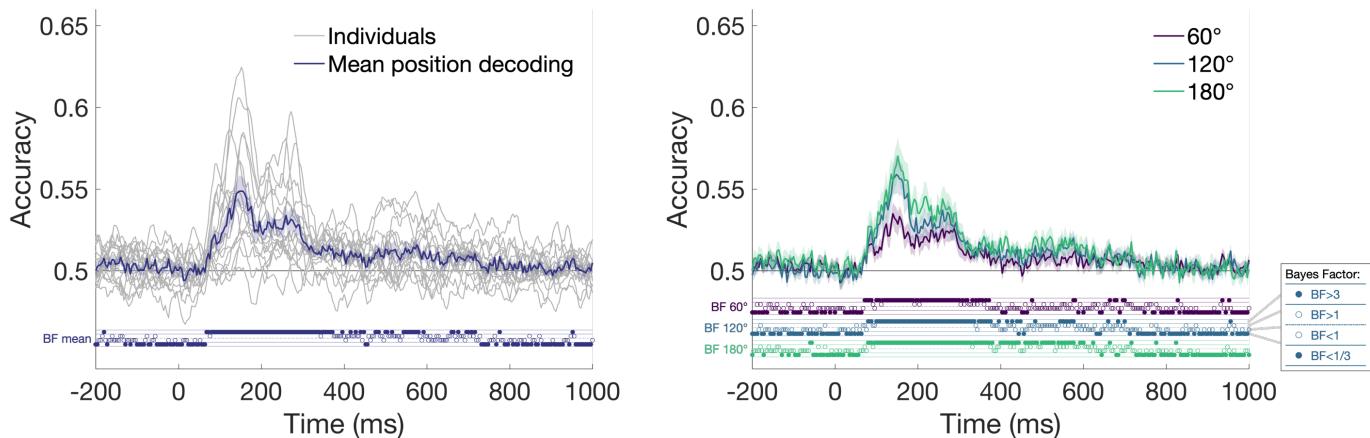
238 Figure 2. Behavioural results. a) Accuracy, and b) Reaction time on the tracking task as a function of final probe position.  
239 Individual subject data are plotted in grey, with group mean in navy. Error bars depict one standard error of the mean across  
240 participants ( $N = 16$ ).

241

242 *Position decoding using the template pattern estimator sequences*

243 The template pattern estimator sequences were designed to extract position-specific neural patterns of  
244 activity from unpredictable visible stimuli. Time-resolved multivariate pattern analysis (MVPA) was  
245 applied to the EEG data from the pattern estimator, which revealed that stimulus position could be  
246 decoded above chance from approximately 68ms after stimulus onset and peaked at 150ms (Figure 3),  
247 consistent with initial retinotopic processing of position in early visual areas (Di Russo et al., 2003; Hagler  
248 et al., 2009). To assess how the physical distance between stimulus positions influenced the neural  
249 patterns of activity, we compared the pairwise decodability of position according to the relative angle

250 between stimulus position pairs (i.e., angle of 60°, 120° or 180° between two stimulus positions). The  
251 greatest decoding performance was observed for larger angles between stimulus positions.



252  
253 Figure 3. Position decoding using template pattern estimator sequences. Left plot shows group mean decoding and smoothed  
254 individual participant decoding for all pairs of positions, and right plot shows mean position decoding as a function of the  
255 angular distance between stimulus pairs. Shaded areas show standard error across participants ( $N = 16$ ). Thresholded Bayes  
256 factors (BF) for above-chance decoding are displayed above the x-axes for every time point as an open or closed circle in one  
257 of four locations (see inset).

258

### 259 *Position decoding on the tracking task*

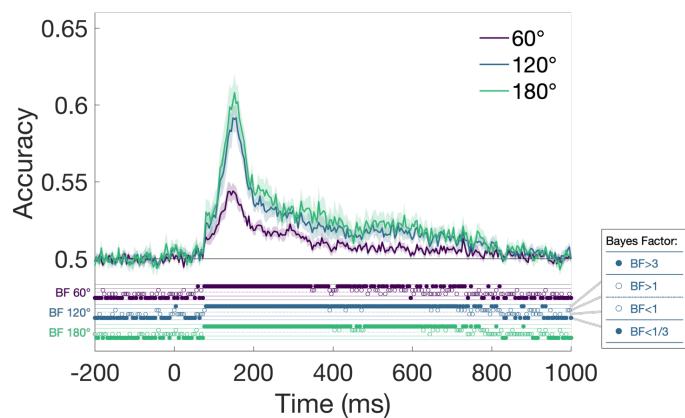
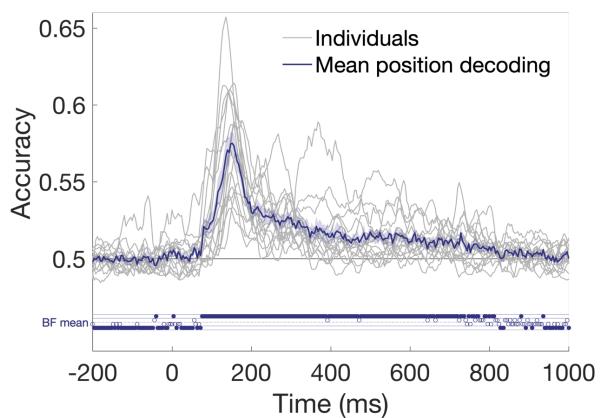
260 To assess the similarity in position representations for visible and invisible (simulated occluded) stimuli,  
261 the classifier was trained on data from the visible template pattern estimator stimuli and tested on data  
262 from the tracking task for the visible and invisible stimuli. Crucially, position could be decoded for both  
263 visible and invisible stimuli, suggesting that similar neural processes underpin perceptual and internal  
264 representations of stimulus position. For visible stimuli, the pattern of decoding results echoed those of  
265 the pattern estimator, with decoding evident from approximately 76ms and peaking at 152ms,  
266 presumably reflecting visual coding of position in ventral visual areas of the brain (Figure 4a, left). When  
267 decoding was split according to the distance between the pair of positions, results looked similar to the  
268 pattern estimator results (Figure 4a, right).

269

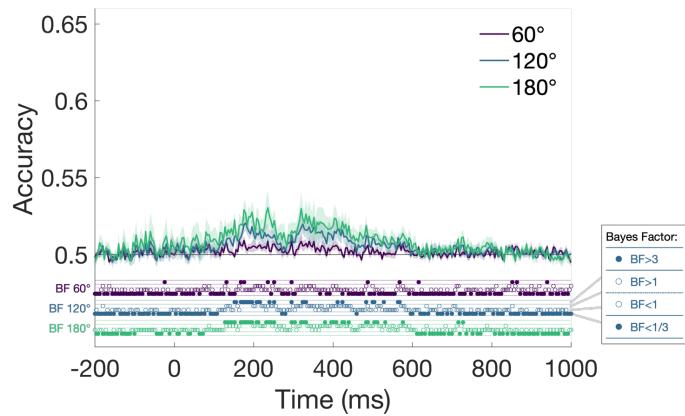
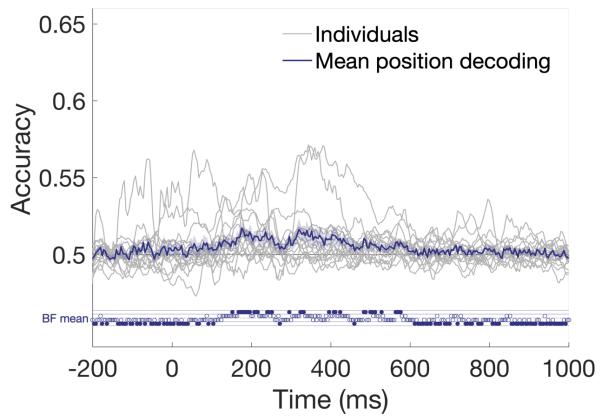
270 A different pattern of results was observed for the invisible stimuli. Here, decoding was not above chance  
271 until approximately 152ms and peaked at 176ms (Figure 4b). The above chance cross-decoding from the  
272 visible pattern estimator stimuli to the invisible stimuli on the tracking task indicates that overlapping  
273 processes underlie stimulus-driven and internally-generated representations of spatial location. But this  
274 decoding of the internal representation of position was later and less accurate than position decoding for  
275 visible stimuli. Similar to the pattern estimator and visible decoding results, positions that were further  
276 apart were more decodable (Figure 4b, right). Notably, neighbouring positions ( $60^\circ$  apart) showed little  
277 evidence of position decoding, suggesting that the representations of position were spatially diffuse for  
278 the invisible stimuli, unlike for the visible stimuli.

279

### A. Visible Stimuli



### B. Invisible Stimuli



280

281 Figure 4. Position decoding from object tracking task. a) Visible stimuli. b) Invisible stimuli. Left plots show group mean decoding  
282 and smoothed individual participant decoding for all pairs of positions, and right plots show mean position decoding as a  
283 function of the angular distance between position pairs. Shaded areas show standard error across participants ( $N = 16$ ).

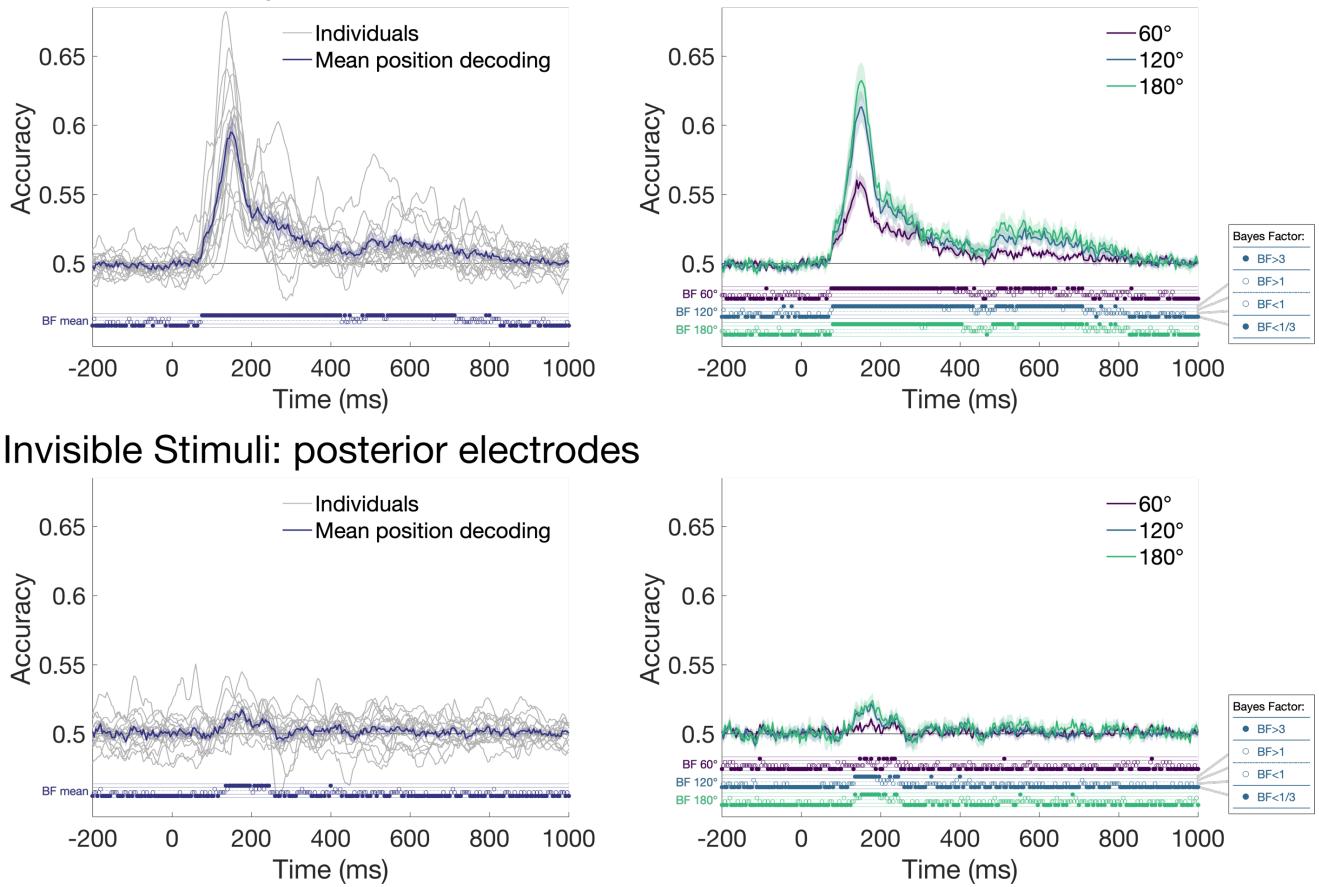
284 Thresholded Bayes factors (BF) for above-chance decoding are displayed above the x-axes for every time point as an open or  
285 closed circle in one of four locations (see inset).

286

287 The previous analyses were performed using electrodes covering the whole head, which meant that there  
288 was a possibility that non-neural artefacts such as eye movements might contribute to the classification  
289 results (Quax et al., 2019). Saccadic artefacts tend to be localised to frontal electrodes, close to the eyes  
290 (Lins et al., 1993). To assess if the EEG signal contributing to the position-specific neural information  
291 originated from posterior regions of the brain (e.g., occipital cortex), as expected, we conducted the same  
292 time-resolved decoding analyses using a subset of electrodes from the back half of the head. We used 28  
293 electrodes that were likely to pick up the largest signal from occipital, temporal and parietal areas (and  
294 were less likely to be contaminated with frontal or muscular activity). The electrodes were CPz, CP1, CP2,  
295 CP3, CP4, CP5, CP6, Pz, P1, P2, P3, P4, P5, P6, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, O2, TP7, TP8, TP9  
296 and TP10. As can be seen in Figure 5, the same trend of results was seen using this subset of electrodes  
297 compared with the whole head analyses in Figure 4. Specifically, Bayes Factors revealed evidence that  
298 position of invisible stimuli was decodable approximately 136-244 ms, which is slightly earlier than the  
299 whole brain results. Decoding was also most evident for positions that were a distance of 120° or 180°  
300 apart (Figure 5b). Analyses restricted to frontal electrodes showed later, more diffuse coding for visible  
301 stimuli, and little evidence for position coding of invisible stimuli (see Figure S2, <https://osf.io/8v47t/>).  
302 Thus, position-specific neural information for visible and invisible stimuli was evident specifically over  
303 posterior regions of the brain, consistent with visual cortex representing stimulus-driven and internal  
304 representations of spatial location.

305

## A. Visible Stimuli: posterior electrodes



306

307 Figure 5. Position decoding from object tracking task using only posterior electrodes. a) Visible stimuli. b) Invisible stimuli. Left  
308 plots show group mean decoding and smoothed individual participant decoding for all pairs of positions, and right plots show  
309 mean position decoding as a function of the angular distance between stimulus pairs. Shaded areas show standard error across  
310 participants ( $N = 16$ ). Thresholded Bayes factors (BF) for above-chance decoding are displayed above the x-axes for every time  
311 point as an open or closed circle in one of four locations (see inset).

312

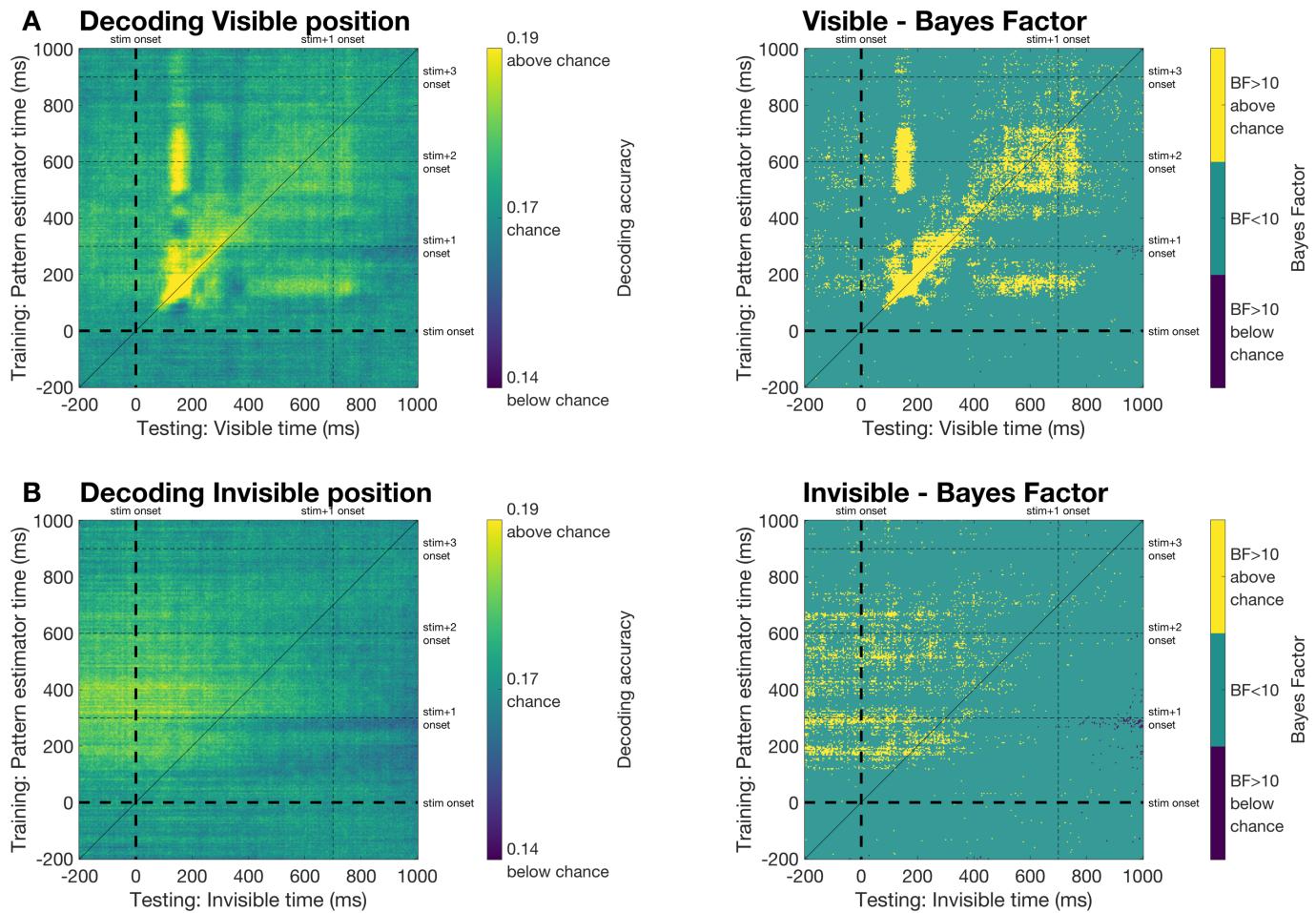
313 The results of the time-resolved analyses showed that position-specific neural patterns for visible stimuli  
314 generalised to invisible stimuli, but with different temporal dynamics. To assess the possibility that neural  
315 processes were more temporally variable for invisible than for visible stimuli, we performed whole brain  
316 (63-channel) time-generalisation analyses by training the classifier on all time points of the pattern  
317 estimator and testing on all time points from the tracking task. As expected, position could be decoded  
318 from both visible and invisible stimulus presentations, but with marked differences in their dynamics  
319 (Figure 6). For the visible stimuli, most of the above-chance decoding was symmetric on the diagonal,

320 indicating that the position-specific processes occurred at approximately the same time for visible stimuli  
321 in the pattern localiser and the tracking task (Figure 6a), even though the inter-stimulus intervals for  
322 stimuli in the training and test sets were different. Interestingly, there was also some above-diagonal  
323 decoding indicating that some neural signals observed in the pattern localiser occurred substantially  
324 *earlier* in the tracking task, which may reflect prediction based on the previous stimuli. Also likely  
325 reflecting anticipation of the stimulus position, generalisation occurred for time points prior to onset of  
326 the visible stimulus in the tracking task. After the tracking stimulus was presented (800-1000ms), there is  
327 some evidence of below chance decoding, indicating a different stimulus position was systematically  
328 predicted. This is likely to reflect processing of the next stimulus in the tracking task, which was presented  
329 at 700ms on the plot (stim +1 vertical line).

330

331 Time generalisation for the invisible stimulus position was not centred on the diagonal, reflecting different  
332 temporal dynamics for the predicted internal representations than for the stimulus-driven processing of  
333 the template pattern estimator. Decoding generalisation was also much more diffuse and relied on  
334 processes approximately 120-750 ms after stimulus onset in the pattern estimator (Figure 6b). Decoding  
335 again preceded the onset of the tone in the tracking task, reflecting an anticipation effect. There was also  
336 below chance decoding at later time points, indicating that the classifier was predicting a different  
337 stimulus position at times when the next stimulus would be processed. Overall, time generalisation results  
338 suggest that during the invisible stimulus portion of the tracking task, which relied on internal  
339 representations of position, the neural dynamics were more variable and anticipatory.

340



341

342 Figure 6. Time generalisation results. a) Decoding visible stimulus position. b) Decoding invisible stimulus position. Decoding  
343 was performed by training on data from the template pattern estimator sequences of visible stimuli and testing on the  
344 experimental trials for all pairs of time points. Left plots show 6-way decoding accuracy for stimulus position, and right plots  
345 show associated Bayes Factors.

346

## 347 Discussion

348 In this study, we assessed the neural underpinnings of internally-generated representations of spatial  
349 location. Participants viewed predictable sequences of a moving stimulus and imagined the sequence  
350 continuing when the stimulus disappeared. Time-resolved MVPA revealed that patterns of activity  
351 associated with visual processing in random sequences were also associated with processing of visible  
352 and invisible spatial stimulus positions in the tracking task, but with different temporal dynamics.  
353 Specifically, the neural correlates of invisible position (i.e., internally-generated representations) were  
354 anticipatory and more temporally diffuse than those of visible position (i.e., sensory and perceptual

355 representations). Taken together, this study provides evidence that internal representations of spatial  
356 position rely on mechanisms of visual processing, but that these are applied with different temporal  
357 dynamics to actual perceptual processes.

358

359 The results of this study suggest that similar perceptual processes are implemented for processing  
360 position of visible and invisible (e.g., occluded) stimuli. This adds to previous neuroimaging work using  
361 high level objects by showing that internally-generated spatial representations appear to use the same  
362 visual perceptual processes as viewed stimuli (Dijkstra et al., 2018). What neural processes are responsible  
363 for this low-level spatial imagery? We found generalisation from the template pattern estimator to the  
364 visible tracked stimuli began at approximately 76ms, but for invisible stimuli the generalisation did not  
365 occur until 120ms. This suggests that internal spatial representations do not rely on early retinotopic  
366 processes such as that of V1, but are implemented by higher order visual processes. Above-chance  
367 generalization for visible and invisible stimuli was maintained until approximately 750ms after the pattern  
368 estimator stimulus was presented, indicating that position-specific information represented throughout  
369 the visual hierarchy has some similarity for stimulus-driven and internally generated representations. It is  
370 important to note, however, that the time generalisation results did not show evidence of distinct,  
371 progressive stages of processing for the invisible representations. In contrast, the visible stimuli showed  
372 different clusters of above-chance decoding on the diagonal of the time-generalisation results, indicating  
373 that there were distinct stages of processing. These results are similar to those observed in Dijkstra et al.,  
374 (2018) during imagery of faces and houses. Internal representations thus seem to activate different  
375 perceptual processes simultaneously, rather than the representations involving information flow through  
376 different brain regions.

377

378 For both visible and imagined stimuli, more distant stimulus positions could more easily be discriminated  
379 by the EEG signals. Decoding for neighbouring positions ( $60^\circ$  apart) was generally much lower than

380 decoding for positions that were further apart. This is consistent with the retinotopic organization of visual  
381 cortices (Tootell et al., 1998), where closer areas of space are represented in neighbouring regions of  
382 cortex, leading to more similar spatial patterns of activation that are measured on the scalp with EEG  
383 (Carlson et al., 2011). Time generalization results also showed that neural patterns of activity from the  
384 template pattern estimator sequences generalized above chance to neighbouring positions. Interestingly,  
385 however, decoding for the closest positions was particularly low for the invisible stimuli, raising the  
386 possibility that internally generated representations of position are more spatially diffuse than perceptual  
387 representations. Together, increasing decodability of stimulus position with increasing distance between  
388 stimuli supports a common, retinotopic mechanism for processing position of both visible and imagined  
389 stimuli, but with greater precision for visible stimuli.

390

391 Another cognitive process that might contribute to the extracted position-specific signal in the current  
392 study is that of spatial attention. In our experimental task, participants were explicitly asked to track the  
393 position of the stimulus, and they performed well, suggesting they were directing their attention to the  
394 location of the stimulus. Spatial attention influences the amplitude of early EEG responses (for review,  
395 see Mangun, 1995), and MEG classification work has shown that spatial attention enhances object  
396 decoding at early stages of processing (Goddard et al., 2019). It is important to note, however, that our  
397 classification results were obtained from training on the template pattern estimator, in which there was  
398 no explicit task and therefore no incentive to specifically attend to stimulus position. The neural patterns  
399 of activity associated with position were therefore more likely to be associated with perceptual rather  
400 than attentional mechanisms. A role of spatial attention cannot be ruled out, however. In the pattern  
401 estimator there was only one stimulus presented at a time and the onsets were likely to attract attention,  
402 albeit in a different fashion to the cued positions in the experimental tracking task. It is possible that a  
403 combination of both perceptual and attentional mechanisms is necessary for the generation of internal  
404 spatial representations. Future work could attempt to disentangle the role of perceptual and attentional

405 processes in spatial imagery with a manipulation to reduce attention during the pattern estimator or even  
406 make the stimuli invisible.

407

408 One factor that we tried to control in this study was eye movements. Recent work has shown that even  
409 when participants were instructed to maintain central fixation, the spatial position of a peripheral  
410 stimulus could be decoded from eye movements, and the eye movements appeared to account for  
411 variance in the MEG signal from 200ms after the stimulus was presented (Quax et al., 2019). To reduce  
412 the likelihood of eye movements influencing our spatial representation results, one countermeasure we  
413 implemented was using independent sequences of randomly-ordered visible stimuli (template pattern  
414 estimator sequences) to extract position-specific patterns from the EEG signal and used these to  
415 generalise to the tracking task. Thus, only neural signals in common between the pattern estimator and  
416 the tracking task could result in above chance decoding. The position sequences in the template pattern  
417 estimator (training set) were randomised, so any incidental eye movements were unlikely to consistently  
418 vary with position. The tracking task implemented both clockwise and counter-clockwise sequences, so if  
419 there were eye movements, across the whole experiment a given position would have two completely  
420 different eye movement patterns. Above-chance cross-decoding from the pattern estimator to the  
421 tracking task was therefore unlikely to be driven by eye movements. Second, all stimuli were presented  
422 briefly (100ms duration), and for a short 200ms inter-stimulus interval during the pattern estimator. This  
423 rapid presentation rate reduced the likelihood that participants would overtly move their eyes, as even  
424 the fastest saccades take at least 100ms to initiate (Fischer & Ramsperger, 1984). Third, we excluded  
425 participants that appeared to move their eyes excessively during the template pattern estimator  
426 sequences, which were the sequences used for training the classifier. Finally, we conducted an additional  
427 analysis using only posterior electrodes to validate that the neural patterns of activity informative for  
428 spatial position were consistent with processes within the visual system (e.g., from occipital cortex).  
429 Decoding from posterior electrodes was similar to the whole-brain results. Furthermore, a similar analysis

430 using only frontal electrodes showed later, more diffuse position decoding for visible stimuli, and  
431 insufficient evidence for position decoding of invisible stimuli (see Figure S2, <https://osf.io/8v47t/>),  
432 indicating that frontal signal or artefacts did not drive decoding of spatial position for visible or imagined  
433 stimuli. Taken together, our finding that spatial position generalised from the pattern estimator to the  
434 tracking task from relatively early stages of processing indicates that it was actually a neural  
435 representation of spatial location that was driving the classifier rather than any overt eye movements.

436

437 In conclusion, in this study we successfully decoded the position of predictable visible and invisible stimuli  
438 using patterns of neural activity extracted from independent visible stimuli. Our findings suggest that  
439 internally generated spatial representations involve mid- and high-level perceptual processes. The visible  
440 stimuli that we used relied on early retinotopic visual processes, yet we found no evidence of  
441 generalisation from very early processes (90-120ms) to the invisible stimuli. The stimuli we used were  
442 much simpler than the vivid, complex objects used in previous work, but we found similar stages of  
443 processing generalised from perceptual to internally-generated representations (Dijkstra et al., 2018),  
444 suggesting a general role of mid- and high-level perceptual processing in internally-generated  
445 representations such as those implemented during imagery or occlusion. Our finding that mid- and high-  
446 level perceptual processes were spatially diffuse and occurred earlier for invisible objects than for the  
447 unpredictable objects indicates an important role of prediction in generating internal representations.  
448 Together, our findings suggest that similar neural mechanisms underlie internal representations and  
449 visual perception, but the timing of these processes is dependent on the predictability of the stimulus.

450

451    **References**

- 452    Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- 453    Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turrey, J. (2011). High temporal resolution  
454        decoding of object position and category. *Journal of Vision*, 11(10), 9–9.  
455        <https://doi.org/10.1167/11.10.9>
- 456    Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG  
457        dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1),  
458        9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- 459    Dentico, D., Cheung, B. L., Chang, J.-Y., Guokas, J., Boly, M., Tononi, G., & Van Veen, B. (2014). Reversal  
460        of cortical information flow during visual imagery as compared to visual perception.  
461        *NeuroImage*, 100, 237–243. <https://doi.org/10.1016/j.neuroimage.2014.05.081>
- 462    Di Russo, F., Martinez, A., & Hillyard, S. A. (2003). Source analysis of event-related cortical activity during  
463        visuo-spatial attention. *Cerebral Cortex*, 13(5), 486–499.
- 464    Dienes, Z. (2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on  
465        Psychological Science*, 6(3), 274–290. <https://doi.org/10.1177/1745691611406920>
- 466    Dijkstra, N., Mostert, P., de Lange, F. P., Bosch, S., & van Gerven, M. A. J. (2018). Differential temporal  
467        dynamics during visual imagery and perception. *eLife*, 7(e33904), 1–16.  
468        <https://doi.org/10.7554/eLife.33904.001>
- 469    Dijkstra, N., Zeidman, P., Ondobaka, S., van Gerven, M. A. J., & Friston, K. (2017). Distinct Top-down and  
470        Bottom-up Brain Connectivity During Visual Perception and Imagery. *Scientific Reports*, 7(1).  
471        <https://doi.org/10.1038/s41598-017-05888-8>
- 472    Fischer, B., & Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal  
473        directed eye movements. *Experimental Brain Research*, 57(1).  
474        <https://doi.org/10.1007/BF00231145>

- 475 Goddard, E., Carlson, T. A., & Woolgar, A. (2019). Spatial and feature-selective attention have  
476 qualitatively different effects on population-level tuning. *BioRxiv*, 530352.  
477 <https://doi.org/10.1101/530352>
- 478 Grootswagers, T., Robinson, A. K., & Carlson, T. A. (2019). The representational dynamics of visual  
479 objects in rapid serial visual processing streams. *NeuroImage*, 188, 668–679.  
480 <https://doi.org/10.1016/j.neuroimage.2018.12.046>
- 481 Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from Evoked  
482 Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging  
483 Data. *Journal of Cognitive Neuroscience*, 29(4), 677–697. [https://doi.org/10.1162/jocn\\_a\\_01068](https://doi.org/10.1162/jocn_a_01068)
- 484 Hagler, D. J., Halgren, E., Martinez, A., Huang, M., Hillyard, S. A., & Dale, A. M. (2009). Source estimates  
485 for MEG/EEG visual evoked responses constrained by multiple, retinotopically-mapped stimulus  
486 locations. *Human Brain Mapping*, 30(4), 1290–1309. <https://doi.org/10.1002/hbm.20597>
- 487 Hogendoorn, H., & Burkitt, A. N. (2018). Predictive coding of visual object position ahead of moving  
488 objects revealed by time-resolved EEG decoding. *NeuroImage*, 171(December 2017), 55–61.  
489 <https://doi.org/10.1016/j.neuroimage.2017.12.063>
- 490 Ishai, A. (2002). Visual Imagery of Famous Faces: Effects of Memory and Attention Revealed by fMRI.  
491 *NeuroImage*, 17(4), 1729–1741. <https://doi.org/10.1006/nimg.2002.1330>
- 492 Jeffreys, H. (1961). *Theory of probability*. Oxford University Press.
- 493 Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430),  
494 773–795.
- 495 King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal  
496 generalization method. *Trends in Cognitive Sciences*, 18(4), 203–210.  
497 <https://doi.org/10.1016/j.tics.2014.01.002>
- 498 Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., & others. (2007). What's new in  
499 Psychtoolbox-3. *Perception*, 36(14), 1.

- 500 Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Maljkovic, V., Weise, S. B., Chabris, C. F., Hamilton, S. E.,  
501 Rauch, S. L., & Buonanno, F. S. (1993). Visual Mental Imagery Activates Topographically  
502 Organized Visual Cortex: PET Investigations. *Journal of Cognitive Neuroscience*, 5(3), 263–287.  
503 <https://doi.org/10.1162/jocn.1993.5.3.263>
- 504 Kwon, O.-S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception.  
505 *Proceedings of the National Academy of Sciences*, 112(26), 8142–8147.  
506 <https://doi.org/10.1073/pnas.1500361112>
- 507 Le Bihan, D., Turner, R., Zeffiro, T. A., Cuénod, C. A., Jezzard, P., & Bonnerot, V. (1993). Activation of  
508 human primary visual cortex during visual recall: a magnetic resonance imaging study.  
509 *Proceedings of the National Academy of Sciences of the United States of America*, 90(24), 11802–  
510 11805.
- 511 Lee, S. H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world  
512 objects. *NeuroImage*, 59(4), 4064–4073. <https://doi.org/10.1016/j.neuroimage.2011.10.055>
- 513 Lins, O. G., Picton, T. W., Berg, P., & Scherg, M. (1993). Ocular artifacts in EEG and event-related  
514 potentials I: Scalp topography. *Brain Topography*, 6(1), 51–63.  
515 <https://doi.org/10.1007/BF01234127>
- 516 Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. The MIT Press.
- 517 Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, 32(1), 4–18.  
518 <https://doi.org/10.1111/j.1469-8986.1995.tb03400.x>
- 519 Mechelli, A. (2004). Where Bottom-up Meets Top-down: Neuronal Interactions during Perception and  
520 Imagery. *Cerebral Cortex*, 14(11), 1256–1265. <https://doi.org/10.1093/cercor/bhh087>
- 521 Nanay, B. (2010). Perception and imagination: amodal perception as mental imagery. *Philosophical  
522 Studies*, 150(2), 239–254. <https://doi.org/10.1007/s11098-009-9407-5>

- 523 Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and  
524 ERP measurements. *Clinical Neurophysiology*, 112(4), 713–719. [https://doi.org/10.1016/S1388-2457\(00\)00527-7](https://doi.org/10.1016/S1388-2457(00)00527-7)
- 525
- 526 Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern  
527 Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, 10.  
528 <https://doi.org/10.3389/fninf.2016.00027>
- 529 Pearson, J. (2019). The human imagination: the cognitive neuroscience of visual mental imagery. *Nature  
530 Reviews Neuroscience*. <https://doi.org/10.1038/s41583-019-0202-9>
- 531 Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into  
532 movies. *Spatial Vision*, 10(4), 437–442.
- 533 Quax, S. C., Dijkstra, N., van Staveren, M. J., Bosch, S. E., & van Gerven, M. A. J. (2019). Eye movements  
534 explain decodability during perception and cued attention in MEG. *NeuroImage*, 195, 444–453.  
535 <https://doi.org/10.1016/j.neuroimage.2019.03.069>
- 536 Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: Decoding category information  
537 during mental imagery. *NeuroImage*, 50(2), 818–825.  
538 <https://doi.org/10.1016/j.neuroimage.2009.11.084>
- 539 Robinson, A. K., Grootswagers, T., & Carlson, T. A. (2019). The influence of image masking on object  
540 representations during rapid serial visual presentation. *BioRxiv*, 515619.  
541 <https://doi.org/10.1101/515619>
- 542 Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting  
543 and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–237.
- 544 Tootell, R. B., Hadjikhani, N. K., Vanduffel, W., Liu, a K., Mendola, J. D., Sereno, M. I., & Dale, a M.  
545 (1998). Functional analysis of primary visual cortex (V1) in humans. *Proceedings of the National  
546 Academy of Sciences of the United States of America*, 95(3), 811–817.  
547 <https://doi.org/10.1073/pnas.95.3.811>

- 548 Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic  
549 Bulletin & Review*, 14(5), 779–804. <https://doi.org/10.3758/BF03194105>
- 550 Wetzel, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical  
551 Evidence in Experimental Psychology: An Empirical Comparison Using 855 t Tests. *Perspectives  
552 on Psychological Science*, 6(3), 291–298. <https://doi.org/10.1177/1745691611406923>
- 553 Wetzel, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial  
554 correlations. *Psychonomic Bulletin & Review*, 19(6), 1057–1064. [https://doi.org/10.3758/s13423-012-0295-x](https://doi.org/10.3758/s13423-<br/>555 012-0295-x)
- 556 Zellner, A., & Siow, A. (1980). Posterior odds ratios for selected regression hypotheses. In J. M.  
557 Bernardo, M. H. DeGroot, D. V. Lindley, & A. F. M. Smith (Eds.), *Bayesian statistics: Proceedings  
558 of the First International Meeting* (pp. 585–603). University of Valencia Press.
- 559