1 **Dual-format attentional template during preparation in human visual cortex**

2 Yilin Chen[1], Taosheng Liu[2], Ke Jia[3,4,5*], Jan Theeuwes[6], Mengyuan Gong[1*]

3

4 [1] Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, Zhejiang, China

5 [2] Department of Psychology, Michigan State University, East Lansing, Michigan, USA

6 [3] Liangzhu Laboratory, MOE Frontier Science Center for Brain Science and Brain-machine Integration,

7 State Key Laboratory of Brain-machine Intelligence, Zhejiang University, Hangzhou, Zhejiang, China

8 [4] Department of Neurobiology, Affiliated Mental Health Center & Hangzhou Seventh People's Hospital,

9 Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

10 [5] NHC and CAMS Key Laboratory of Medical Neurobiology, Zhejiang University, Hangzhou, Zhejiang,

11 China

12 [6] Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, the

13 Netherlands
14

15 ***Corresponding Author**

16 Dr. Mengyuan Gong

17 Department of Psychology and Behavioral Sciences, Zhejiang University

18 gongmy426@zju.edu.cn

19

20 Dr. Ke Jia

21 MOE Frontier Science Center for Brain Science and Brain-machine Integration, Zhejiang University

22 kjia@zju.edu.cn

23 **Abstract**

24 Goal-directed attention relies on forming internal templates of key information relevant for guiding

25 behavior, particularly when preparing for upcoming sensory inputs. However, evidence on how these

26 attentional templates is represented during preparation remains controversial. Here, we combine

27 functional magnetic resonance imaging (fMRI) with an orientation cueing task to isolate preparatory

28 activity from stimulus-evoked responses. Using multivariate pattern analysis, we found decodable

29 information of the to-be-attended orientation during preparation; yet preparatory activity patterns were

30 different from those evoked when actual orientations were perceived. When perturbing the neural activity

31 by means of a visual impulse ('pinging' technique), the preparatory activity patterns in visual cortex

32 resembled those associated with perceiving these orientations. The observed differential patterns with and

33 without the impulse perturbation suggest a predominantly non-sensory format and a latent, sensory-like

34 format of representation during preparation. Furthermore, the emergence of the sensory-like template

35 coincided with enhanced information connectivity between V1 and frontoparietal areas and was

36 associated with improved behavioral performance. This dual-format mechanism suggests that during

37 preparation the brain encodes more detailed template information beyond its immediate use, potentially

38 providing advantages for adaptive attentional control. Consistent with recent theories of non-veridical,

39 'good-enough' attentional template for initial guidance, our findings established a neural basis for

40 implementing two representational formats in different functional states during preparation: a

41 predominantly non-sensory format for guidance and a latent sensory-like format for prospective stimulus

42 processing.

43

44 **Keywords**

45 Attentional selection, representational format, fMRI, decoding, visual cortex, impulse perturbation

46

Introduction

To address the challenge of processing the overwhelming amount of sensory inputs from the external environment, the brain must allocate attentional resources to prioritize the processing of task-relevant information. Importantly, humans can proactively prepare for stimulus selection before the arrival of sensory inputs (Summerfield & De Lange, 2014) . For example, when preparing to hail a taxi on the road, we tend to form a mental representation of the defining features of a taxi (e.g., yellow with a car-like shape). This ability relies on the formation of attentional templates — mental representations of the target — to accelerate stimulus selection and resolve perceptual competition by enhancing task-relevant information and suppressing irrelevant information (Desimone & Duncan, 1995; Kastner et al., 1999). While most attentional models posit that attentional templates during stimulus processing reflect veridical representations of the target (Jigo et al., 2018; Malcolm & Henderson, 2009), the nature of the template during preparation remains less understood.

A classical view suggests that attentional template during preparation may reflect veridical target features, analogous to the representational format during stimulus selection. However, evidence supporting this account has been mixed. For example, while some previous fMRI studies have demonstrated that preparatory activity contains target information similar to the sensory responses to the corresponding targets (Kok et al., 2014; Lewis-Peacock et al., 2015; Stokes et al., 2009), more recent electrophysiological studies suggest that, if anything, this template is engaged only shortly before the expected arrival of sensory input rather than being continuously active (Grubert & Eimer, 2018; Myers et al., 2015). Notably, in some cases the template is even largely undetectable during preparation (Wen et al., 2019). Alternatively, an emerging view suggests a non-veridical template suffices for guiding attention during preparation, where precise processing of stimuli may be unnecessary at this stage. Support for this notion comes from the identification of attentional signals during preparation that differ from neural signals observed during perceptual target processing (Gong et al., 2022). Recent theories of visual search also propose a non-veridical, "good-enough" template for early attentional guidance (Wolfe, 2021; Yu et

3

72   al., 2023). However, it remains unclear whether a "good-enough" template for search also applies to

73   preparatory attention.

74   The notion that there may be a sensory and non-sensory attentional template might not be as far-

75   fetched as it seems. Indeed, it is feasible that during preparation, following stimulus presentation,

76   attentional signals undergo a transformation from a non-sensory to a sensory-like template. Previous

77   behavioral (Hamblin-Frohman & Becker, 2021; Yu et al., 2022) and neural studies (Gong et al., 2022;

78   Jigo et al., 2018; Wen et al., 2019) are generally consistent with this idea of coarse-to-fine transitions,

79   suggesting that during preparation, a sensory-like template may not be initially necessary but only

80   becomes relevant when the stimulus needs to be identified. However, if and in what way the brain

81   coordinates these non-sensory and sensory-like templates remains unclear. Here, we propose that during

82   preparation, a sensory-like template may be stored in a latent (e.g., activity-silent) state concurrently with

83   a non-sensory template. This idea parallels recent findings from working memory studies, which suggest

84   that information intended for proactive use is kept in activity-silent traces to support future behavior

85   (Stokes, 2015; Wolff et al., 2015, 2017). The present study seeks to determine the possibility of the latent,

86   sensory-like template during the preparation for discriminating an upcoming stimulus.

87   To test these hypotheses, participants engaged in a cuing task in which they prepared during an

88   extended period of time for the presentation of a compound stimulus grating containing the cued

89   orientation and a distractor orientation. In addition, in order to be able to construct the sensory-format

90   representations (leftward and rightward orientation), single orientations were presented during the

91   perception task. Critically, we used a "pinging" technique combined with multivariate decoding methods,

92   which has been shown to be effective in retrieving information from latent brain states (Duncan et al.,

93   2023; Wolff et al., 2015, 2017; Zhang & Luo, 2023). In the standard condition (*No-Ping session*), the

94   preparation period was devoid of visual impulses. During preparation, the neural activity patterns in

95   visual and frontoparietal areas could discriminate between the orientations that participants were

96   preparing for. Yet, neural activity patterns evoked by the perception of orientations were distinct from

97   those evoked during preparation for upcoming orientations, suggesting a predominantly non-sensory

4

98 template during preparation. By contrast, when we presented a high-contrast, task-irrelevant impulse

99 stimulus during preparation (*Ping session*), neural activity patterns evoked by the perception of

100 orientations were similar to those activated by the preparation for orientations in the visual cortex,

101 suggesting the existence of a latent, sensory-like format of representation during preparatory attention.

102 Furthermore, the emergence of sensory-like template coincided with enhanced information connectivity

103 between V1 and frontoparietal areas and was associated with improved behavioral performance. Our

104 findings provide evidence for the co-existence of two formats of attentional templates (non-sensory vs.

105 sensory-like) during preparation, as well as a novel neural mechanism for their maintenance in different

106 functional states (active vs. latent). We propose that this dual-format representation may serve to increase

107 flexibility of attentional control.

108

Results

Behavioral performance during the attention tasks

111 In an orientation cueing task, participants were shown a color cue indicating the reference orientation (45°

112 or 135°) to attend to during preparation period (a delay of 5.5 s or 7.5 s) with or without the impulse

113 perturbation (Fig 1A). This was followed by the presentation of a compound stimulus consisting of two

114 oriented gratings. During the stimulus selection period (after the gratings appeared), participants were

115 tasked with discriminating a small angular offset of the cued grating from the cued reference orientation.

116 The angular offset was individually thresholded before the scanning sessions (mean offset = 2.50° in the

117 No-Ping session and 2.52° in the Ping session) without significant difference between two sessions

118 (independent t-test: $t(19) = 0.085$, $p = 0.932$, Cohen's d = -0.027). Participants' discrimination

119 performance showed no significant difference between two attended orientations in either the No-Ping

120 (paired t-test: $t(19) = 1.439$, $p = 0.166$, Cohen's d = -0.321) or the Ping session (paired t-test: $t(19) =$

121 0.494, $p = 0.627$, Cohen's d = 0.122; Fig 1C). A two-way mixed ANOVA (attended orientation × session)

122 revealed neither significant main effects (attended orientation: $F(1,38) = 0.392$, $p = 0.535$, $\eta_p^2 = 0.01$;

123   session: F(1,38) = 0.001, p = 0.970, η_p^2 < 0.001) nor interaction effect (F(1,38) = 1.811, p = 0.186, η_p^2 =

124   0.045). Bayesian analyses provided moderate evidence to support the null hypothesis (BF_{excl} > 3.633),

125   suggesting comparable performance levels between two sessions and two attended orientations.
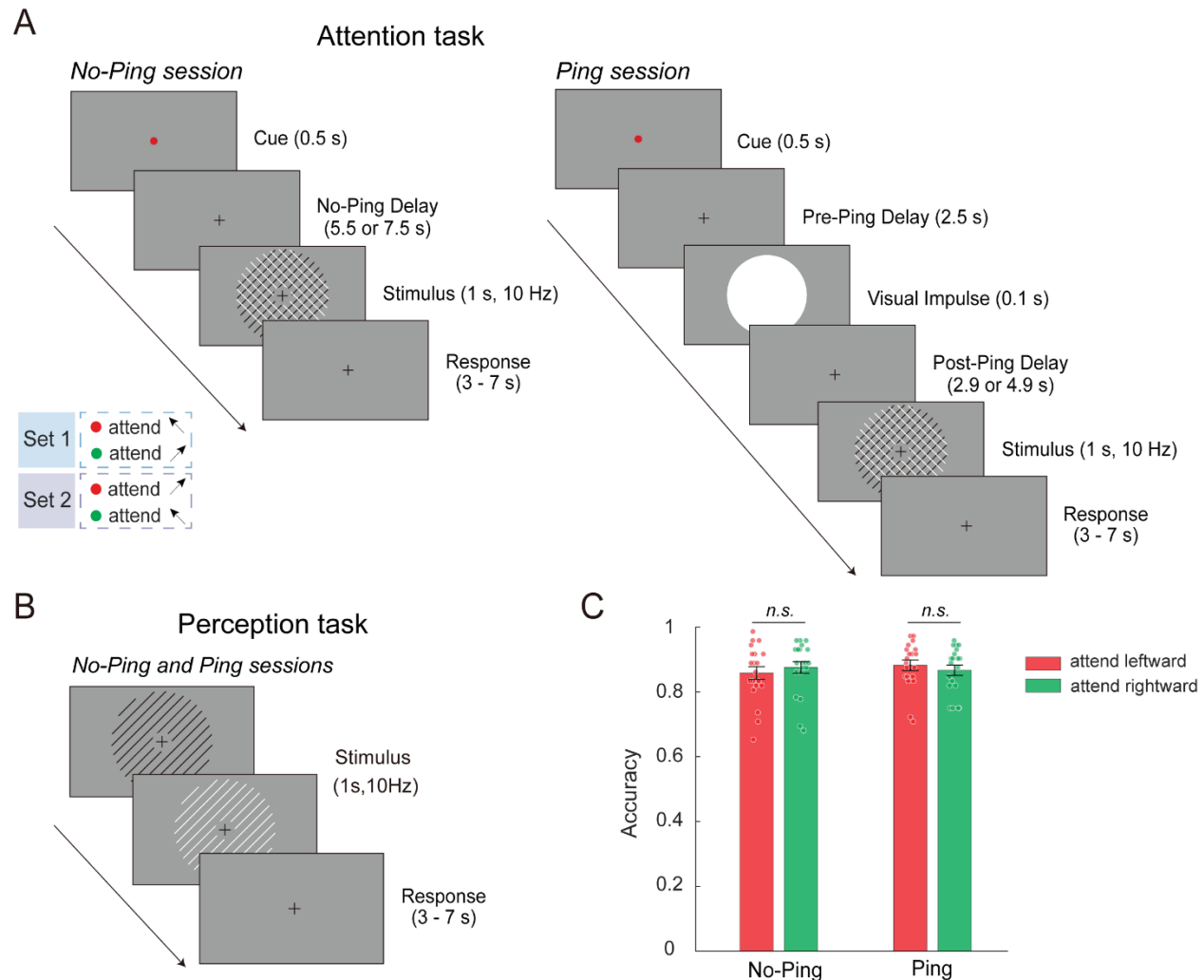
126

127

128   **Fig 1. Experiment procedure and behavioral performance.** (A) Attention tasks. Note that only long-delay trials

129   are shown. A small proportion of short-delay trials (20%, with a delay of 1.5 s or 3.5 s) were included to create

130   temporal uncertainty and encourage consistent active preparation during the delay. Both component gratings were

131   flickering at 10 Hz between white and black, so that luminance could not confound either the task strategy (e.g.,

132   attending to luminance) or neural measures. The inset shows two sets of color-orientation mapping, which were

133   reversed halfway through the experiment to minimize the impact of cue-induced sensory difference on neural

134   activity. (B) Perception task. Similar to the attention task, the single-orientation grating also flickered at 10 Hz

135   between white and black. (C) Behavioral accuracy in the attention tasks in the No-Ping and Ping sessions. Each dot

136   represents one subject's data. Error bars denote standard error of the means (SEM).

137

138

139   **A default, non-sensory representation of attentional template during preparation**

140   The first aim of this study was to determine whether attentional signals during preparation are encoded in

141   a sensory-like or non-sensory format. To address this, we first examined whether, in the attention task

142   during the No-Ping condition, the distributed neural pattern contained feature-specific information. We

143   trained and tested separate classifiers to predict the attended orientation during the preparation and

144   stimulus selection periods (Fig 2A, "Attention decoding"; see Materials and Methods for details). This

145   analysis was performed for each of the four regions along the visual hierarchy, including primary visual

146   cortex (V1), extrastriate visual cortex (EVC), intraparietal sulcus (IPS) and prefrontal cortex (PFC). The

147   average decoding accuracies for both preparation and stimulus selection periods were significantly above

148   chance level in each region (Permutation analyses: $ps < 0.004$ across regions, Fig 2B), indicating that the

149   brain maintained reliable information about the attended feature both before and after the onset of the

150   compound grating. Next, we examined whether the preparatory activity reflected a sensory-like format of

151   attentional template (Fig 2A, "Cross-task generalization", see Materials and Methods). We trained a

152   classifier using data from the perception task (leftward vs. rightward orientation; Fig 1B) and tested its

153   performance on data from the preparation period in the attention task (attend leftward vs. attend

154   rightward). However, this cross-task generalization analysis yielded no significant effects ($ps > 0.132$

155   across the regions). In contrast, we observed above-chance generalization from the perception task to the

156   stimulus selection period ($ps < 0.001$ across regions, Fig 2C), confirming previous findings of the

157   sensory-like attentional template following stimulus presentation (Gong et al., 2022; Jigo et al., 2018;
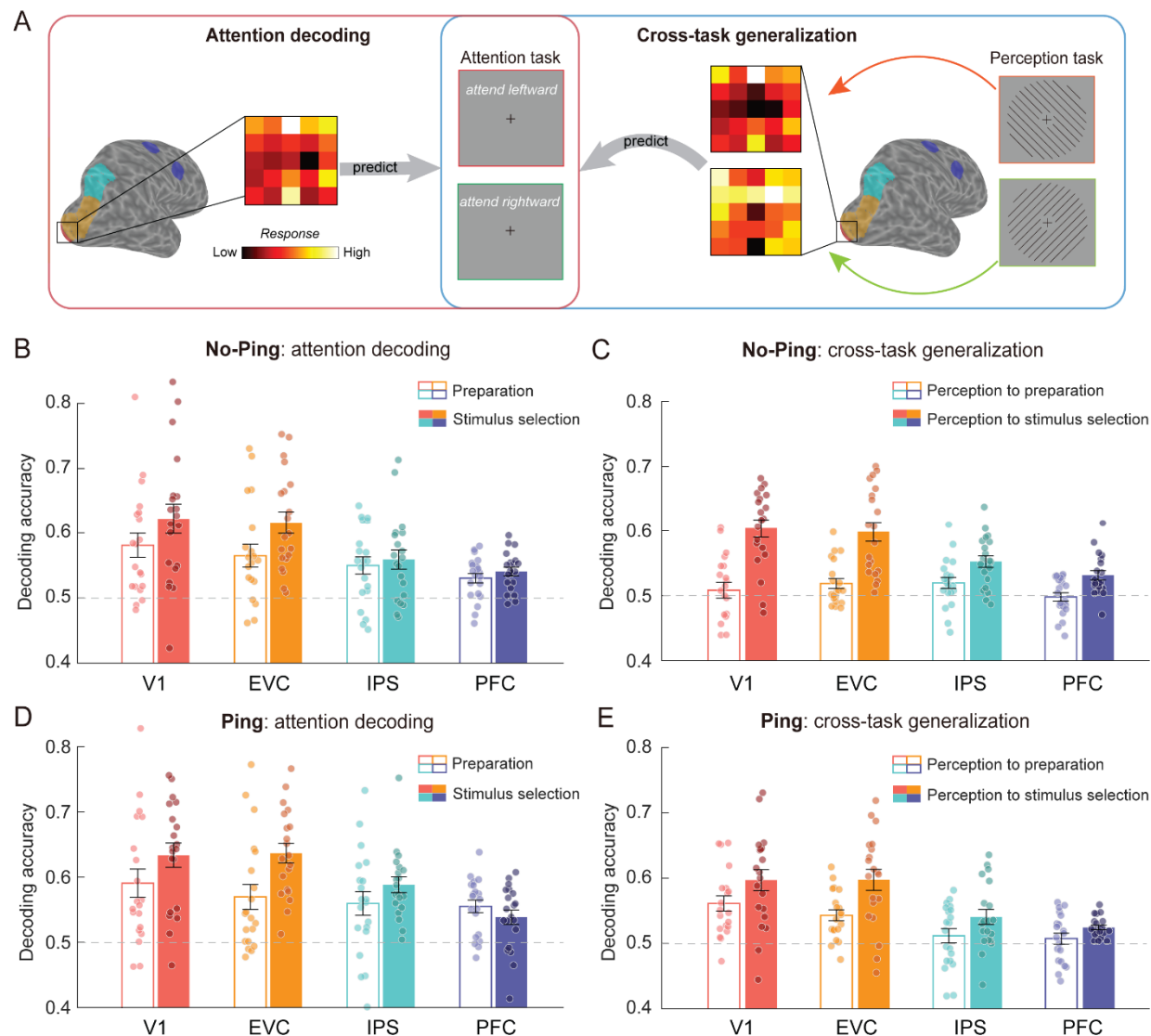
158   Wen et al., 2019).

7

159

Fig 2. MVPA for the No-Ping session and Ping session. (A) Schematic illustration of the decoding of attended orientation (attend leftward vs. attend rightward) in the attention task (left panel) and the cross-task generalization analysis from perception task to the attention task (right panel). The four regions are shown on a representative right hemisphere as colored areas: V1 is marked in red, EVC in yellow, IPS in cyan and PFC in purple. (B) Decoding accuracy during preparation and stimulus selection periods across regions in the No-Ping and (D) Ping session. (C) Cross-task generalization performance from the perception task to the preparatory periods and the stimulus selection periods across regions in the No-Ping and (E) Ping session. The dashed lines represent the theoretical chance level (0.5). Each dot represents one subject's data. Error bars denote SEM.

168

169   Before drawing conclusions based on the lack of generalization from the perception task to

170   preparatory attention, we considered two alternative explanations to rule out potential confounds. First,

171   the robust attention decoding during preparation ruled out the possibility that participants were not

172   actively engaged in the task during preparation (Fig 2B, unfilled bars). Second, the generalizable effect

173   from the perception task to the stimulus selection period across regions (Fig 2B, filled bars) argues

174   against the possibility of low statistical power. Overall, these findings suggest that the preparatory

175   attention and sensory processing of features have distinct formats, presumably reflecting a non-sensory

176   format of representation during the preparation. These results replicate those of a previous fMRI study

177   using motion stimuli with a similar design(Gong et al., 2022). Furthermore, consistent with previous

178   studies(Gong et al., 2022; Jigo et al., 2018), univariate analysis did not reveal any reliable difference in

179   overall BOLD responses between attention orientations (Supplementary Information and S1 Fig).

180

181   **A latent, sensory-like attentional template during preparation revealed by visual impulse**

182   The second aim of our study was to examine whether a latent, sensory-like template exists during

183   preparation. While this precise template may not be necessary for preparation, it is relevant for

184   subsequent target selection and discrimination (i.e., select the cued grating from the compound stimulus

185   and discriminate a small angular offset between the cued grating and the reference orientation). To test

186   this hypothesis, we perturbed the neural activity by means of a visual impulse during the preparation

187   period in the Ping session (Fig 1A, right panel). Using the same analyses as those performed in the No-

188   Ping session, robust attentional signals were observed during both preparation and stimulus selection

189   periods (Permutation analyses: $ps < 0.001$ across regions; Fig 2D). Importantly, the cross-task

190   generalization analyses indicated that the visual impulse led to above-chance generalization from the

191   perception task to preparation period (Fig 2E, unfilled bars) in V1 and EVC ($ps < 0.001$), but not in IPS

192   and PFC ($ps > 0.584$), along with generalizable effects from the perception task to the stimulus selection

193   periods ($ps < 0.036$ across regions; Fig 2E; filled bars). These results suggest that different brain areas are

194   involved in coding for sensory-like templates. To further evaluate whether the cross-task generalization

9

195 from the perception task to the preparation period was statistically different with and without visual

196 impulse, we conducted a two-way mixed ANOVA (session × region) on the generalization performance.

197 The analysis revealed main effects of region (F(3,114) = 5.220, p = 0.002, η_p^2 = 0.121), session (F(1,38) =

198 7.321, p = 0.010, η_p^2 = 0.162), and importantly, a significant interaction effect (F(3,114) = 3.964, p =

199 0.010, η_p^2 = 0.094), supporting the observation that the visual impulse led to significantly increased

200 decoding accuracy in V1 (independent t-test: t(38) = 3.145, p = 0.003; Cohen's d = 0.995) and EVC

201 (independent t-test: t(38) = 2.153, p = 0.038, Cohen's d = 0.681), but not in the frontoparietal regions (*ps* >

202 0.374). This dissociable result between the two sessions further supports the activation of a latent,

203 sensory-like template by the visual impulse during preparatory attention.

204   To further solidify this conclusion, following analyses were used to examine several alternative

205 possibilities. First, we examined whether the impulse-driven generalization resulted from stronger feature

206 information in the Ping compared to No-Ping session during the perception task. This was not the case, as

207 evidenced by comparable levels of decodable orientation information between the Ping and No-Ping

208 sessions (see Supplementary Information and S2 Fig). Next, we asked whether the increased

209 generalization was due to generally stronger attentional signals in the Ping session during the attention

210 tasks. This was not the case, as the two-way mixed ANOVAs (session × region) on attention decoding

211 accuracy revealed neither a significant main effect of session nor an interaction effect during both the

212 preparation (*ps* > 0.519; BF_{excl} > 3.247) and stimulus selection periods (*ps* > 0.336; BF_{excl} > 3.297),

213 suggesting comparable amount of attentional information between the two sessions. Therefore, the

214 findings of impulse-driven sensory-like template in the visual cortex during preparation cannot be

215 explained by general differences between two sessions.

216

217 **Matching preparatory attention to sensory template: impact on neural representation and**

218 **behavior**

219 The reported decoding accuracy from the cross-task generalization analysis quantifies the degree to which

220 differences in neural activity pattern between two conditions are shared across attention and perception

221   tasks. However, it does not directly measure how similar the neural patterns are when attending to an

222   orientation compared to perceiving that orientation. Unlike decoding accuracies, Mahalanobis distance

223   provides a continuous measure for characterizing representational geometries between different

224   conditions (Mahalanobis, 1936). To further corroborate our findings of the impulse-driven sensory-like

225   template, we calculated the Mahalanobis distance between each attention condition during preparation

226   and each perception condition (see Materials and Methods). If the patterns of activity reflect a sensory-

227   like template, we would expect greater pattern similarity (smaller distance) between "attend leftward" and

228   "perceive leftward" than between "attend leftward" and "perceive rightward", and vice versa for the

229   "attend rightward" conditions (see Fig 3A for a schematic of the four pair-wise distance measures),

230   leading to an interaction between attended and perceived orientation conditions.

231   We used a two-way repeated-measures ANOVA (attended orientation × perceived orientation) on

232   the Mahalanobis distance, separately for each session and each region. During the preparatory period in

233   the No-Ping session, no significant interaction effects were observed across regions ($ps > 0.443$; Fig 3B).

234   In contrast, the same analyses applied to the Ping session revealed significant interaction effects in visual

235   areas (V1: $F(1,19) = 9.335$, $p = 0.007$, $\eta_p^2 = 0.329$; EVC: $F(1,19) = 8.563$, $p = 0.009$, $\eta_p^2 = 0.311$; Fig 3C),

236   but not for frontoparietal regions ($ps > 0.213$). This cross-region difference is consistent with the function

237   of sensory areas in encoding precise neural representations for basic visual features. To directly compare

238   whether the attentional modulation on Mahalanobis distance was statistically different with and without

239   the visual impulse, we used a three-way mixed ANOVA (session × attended orientation × perceived

240   orientation). The analysis revealed a significant three-way interaction in V1 ($F(1,38) = 5.00$, $p = 0.031$,

241   $\eta_p^2 = 0.116$), suggesting that by "pinging" the brain, the attentional template during preparation became

242   more similar to the perception of corresponding orientation. We also calculated the Mahalanobis distance

243   between neural patterns evoked by the superimposed gratings during the stimulus selection period and

244   each condition in the perception task, finding similar results (see Supplementary Information and S3 Fig).

245   This result was expected, as feature-based attention is known to selectively enhance the representation of

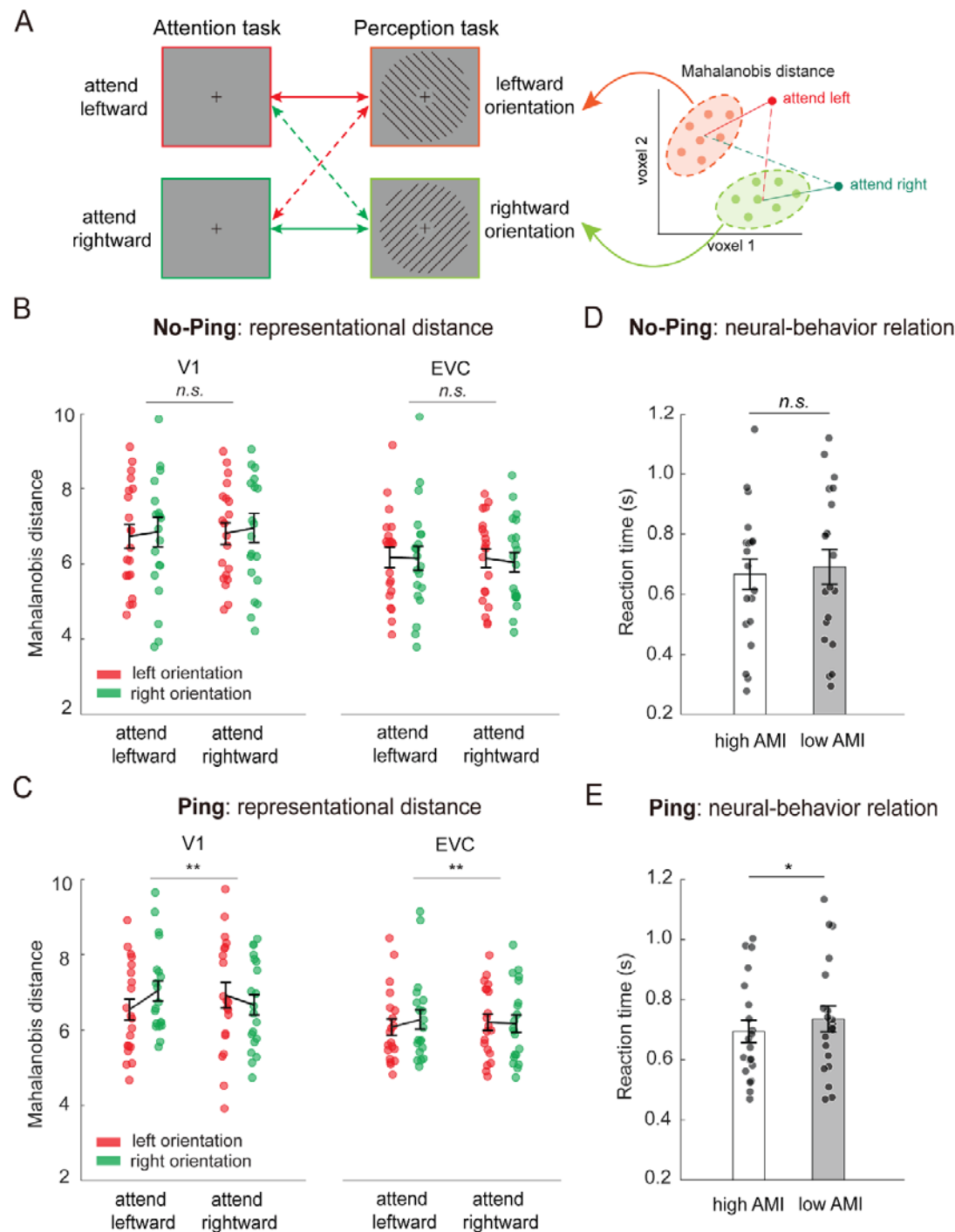246   task-relevant features while filtering out task-irrelevant ones.

Fig 3. Orientation-selective attentional modulations on neural pattern distances during preparation. (A) Schematic illustration of the representational distance (mean Mahalanobis distances) between each of the attention conditions and each of the perception conditions. Colored arrows indicate measures of the pair-wise Mahalanobis distance. The right panel shows two attention trials (red indicates attend-to-leftward and green indicates attend-to-rightward) to the distribution of each perception condition (shown in a cloud of light-colored dots). (B) Mahalanobis

253　distance between preparatory attention condition and perceived orientation condition in the No-Ping and (C) Ping

254　sessions. (D) Behavioral performance (as indexed by reaction time) for "strong modulation" (high AMI) trials and

255　"weak modulation" (low AMI) trials, sorted by V1 preparatory activity in the No-Ping and (E) Ping sessions. Each

256　dot represents one subject's data. Error bars denote SEM. * $p < 0.05$, ** $p < 0.01$.

257

258　　　　　The continuous nature of the Mahalanobis distance also made it possible to further investigate

259　potential neural-behavioral correlations. We examined whether activating a sensory-like template during

260　preparation would benefit subsequent orientation processing. In particular, we calculated attentional

261　modulation indices (AMIs) based on trialwise Mahalanobis distance in V1. The index was calculated as

262　follows: $AMI = (D_{different} - D_{same})/(D_{different} + D_{same})$, where D_{same} and $D_{different}$ are the measured distance (D)

263　in the Same (e.g., attend and perceive the same orientation) and Different (e.g., attend and perceive

264　different orientations) orientation condition, respectively (see Methods and Materials). We then sorted the

265　AMI values in a descending order and selected the top-ranked 25% of trials (i.e., 36 trials) and bottom-

266　ranked 25% of trials to represent "strong modulation" and "weak modulation" trials, respectively and

267　calculated reaction time and accuracy for each group of trials. The analysis revealed a faster response in

268　the "strong modulation" than "weak modulation" trials in Ping session (paired t-test: $t(19) = -2.746$, $p = $

269　0.013, Cohen's d = -0.614; Fig 3E), but not in the No-Ping session (paired t-test: $t(19) = -1.487$, $p = 0.154$,

270　Cohen's d = -0.332; Fig 3D). These effects were not observed for accuracy ($ps = 0.163$). These results

271　suggest that the impulse-driven sensory-like template in primary visual cortex is functionally relevant to

272　subsequent attentional selection, providing evidence for the prospective use of sensory-like template in

273　this task. In addition, we did not observe such behavioral differences in analogous analyses using data

274　from the stimulus selection period in either session ($ps > 0.230$), which might due to the potential dilution

275　by strong stimulus-evoked responses during the stimulus selection period.

276

277　**Activating sensory-like template strengthens the informational connectivity between**

278　**sensory and frontoparietal areas**

13

279   Selective attention is generally believed to rely on coordinated network activity (Corbetta & Shulman,

280   2002). In particular, studies have shown that functional connectivity between sensory and frontoparietal

281   areas was modulated by attentional control (Bressler et al., 2008; Rosenberg et al., 2020). Given that the

282   impulse-driven sensory-like template facilitated behavior, we reasoned that it may also enhance network

283   communication. Thus, we examined informational connectivity measures to explore how the impulse

284   altered network function during the attention task. We used a method that allows inference based on

285   multivoxel pattern information rather than univariate BOLD response (Jia et al., 2020; Ng et al., 2021).

286   For each ROI, we calculated the cross-validated Mahalanobis distance from each attention trial (from one

287   left-out run) to the distribution of each attended orientation (all trials from remaining runs) during

288   preparation (see Methods and Materials). To quantify the degree of attentional modulation during

289   preparation, we calculated the AMI based on trialwise Mahalanobis distance and generated a time course

290   of AMI values across trials. Pearson correlation was used to estimate the covariation between each pair of

291   ROIs, the resulting correlation coefficients were transformed using Fisher's z-transform for statistical

292   inference (Fig 4A). The analysis revealed numerically higher levels of connectivity in Ping than in No-

293   Ping session, this impulse-driven increase in connectivity reached statistical significance in two pairs (Fig

294   4B): V1-IPS (independent t-test: $t(38) = 2.566$, $p = 0.014$; Cohen's $d = 0.812$) and V1-PFC (independent

295   t-tests: $t(38) = 3.158$, $p = 0.003$; Cohen's $d = 0.999$). The enhanced functional connectivity between V1

296   and frontoparietal areas driven by the impulse may potentially facilitate information flow among areas to

297   improve attentional control. Additionally, the same analysis of AMI based on cross-validated

298   Mahalanobis distance during the stimulus selection period showed no significant differences in

299   information connectivity between No-Ping and Ping sessions ($ps > 0.224$; Supplementary Information and

300   S4A Fig). The lack of changes in long-range connectivity during stimulus selection period may be

301   attributed to a general rise in connectivity caused by strong sensory inputs in this period, which could

302   have attenuated any potential impacts of visual impulses. Furthermore, connectivity analysis based on

303   mean BOLD response over time did not reveal any significant changes in inter-cortical connections

304   between the two sessions ($ps > 0.136$; S4B Fig), suggesting that the impulse-driven increased information

305   connectivity between V1 and higher-order areas was unlikely contributed by the overall changes of
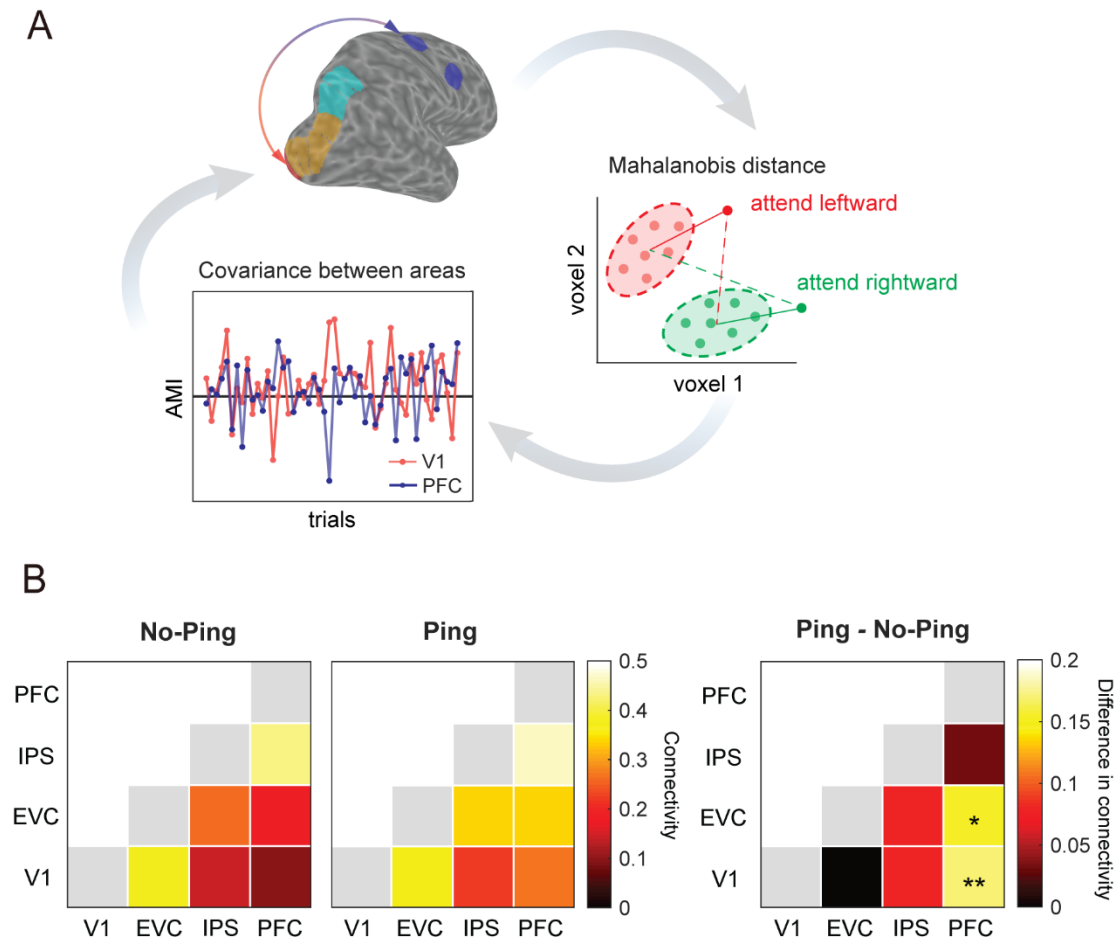
306   BOLD response.

307

Fig 4. Information connectivity analysis. (A) Schematic illustration of the procedure for the information

connectivity analysis in the space of two hypothetical voxels. For each region, we calculated the Mahalanobis

distance of the attention trial (from one left-out run) from two attention distributions (all trials from remaining runs).

Red and green dots indicate activity patterns from two trials (right panel). The brain image shows an example pair of

intercortical information connectivity between V1 and PFC. The time series (lower-left panel) consisted of

attentional modulation index (AMI) based on the Mahalanobis distance. (B) Between-region information

connectivity in the No-Ping session and Ping session (left panel), and differences in connectivity between the two

sessions (right panel). $* \, p < 0.05$, $** \, p < 0.01$.

317

Discussion

While there is ample evidence that the brain can maintain an attentional template of an upcoming target before sensory information is presented (Desimone & Duncan, 1995; Kastner et al., 1999; Summerfield & De Lange, 2014), its representational format remains unclear. To address this, we used an orientation cueing paradigm with separated preparation and stimulus selection periods and applied MVPA to decode neural activity patterns associated with feature-specific attentional information of the upcoming target. The analyses showed robust attentional information both before and after the presentation of the compound grating, suggesting sustained maintenance of attentional templates throughout a trial. Importantly, while the decoders trained on the perception of single orientations could not generalize to preparation until the stimulus selection period (*No-Ping session*), perturbing the brain with a visual impulse resulted in generalizable activity patterns during preparation in the visual cortex (*Ping session*). These results suggest a predominantly non-sensory format of representation, with a sensory-like template in a latent state during feature-based preparation in the visual cortex. Furthermore, impulse-driven sensory-like template was accompanied by enhanced information connectivity between V1 and frontoparietal areas, as well as enhanced orientation-specific neural modulations of neural distances in the visual areas that predicted levels of behavioral performance. The differences between the Ping and No-Ping sessions could not be attributed to differences in sensory information from the perception task, overall strength of preparatory attention, or differences in eye position (S5 Fig). Therefore, our findings suggest dual-format neural representations (non-sensory vs. sensory-like) operating in different functional states (active vs. latent). This mechanism may give rise to flexible attentional control, allowing effective transition from coarse to fine attentional templates at various processing stages (initial guidance vs. precise discrimination).

Recent advances in theories of visual search differentiated between the "guiding template" and "target template" based on measures of behavioral performance and eye movements (Wolfe, 2021; Yu et al., 2023). According to these theories, early attentional guidance typically depends on a non-veridical

16

343   template that includes only the most diagnostic information, whereas later, target-match processes utilize

344   more precise codes to optimize decision accuracy (Kerzel, 2019; Scolari et al., 2012; Yu et al., 2022). Our

345   study reveals a parallel coding mechanism in the context of feature-based attention, expanding upon these

346   theoretical notions in three key aspects. First, we provide neural evidence for a default, predominantly

347   non-sensory template during preparation, indicating that the concept of a "guiding template", as proposed

348   by visual search theories (Wolfe, 2021; Yu et al., 2023), also applies to preparatory attention in a non-

349   search context. This highlights a shared functional role of a non-veridical attentional template in early

350   guidance across different scenarios. Second, despite the theoretical notion that the brain maintains a more

351   veridical template with detailed target information than is typically utilized to form the guiding template

352   (Wolfe, 2021; Yu et al., 2023), direct evidence supporting this hypothesis is currently lacking. We

353   provide evidence for this notion and propose a plausible neural implementation for preserving a more

354   veridical, sensory-like template in the latent state. A natural question is why the sensory-like template

355   remains latent during preparation. We note that our task requires both coarse and fine featural information.

356   Specifically, during preparation, the brain needs to establish a guiding template that aids in distinguishing

357   the target from the distinct distractor; hence a coarse, non-veridical template suffices. During stimulus

358   selection, the fine discrimination task (reporting the tilted direction of a small angular offset) necessitates

359   the representation of a precise template. It is thus both advantageous and efficient to maintain a sensory-

360   like template in a latent state during preparation, which could facilitate the processing of sensory stimulus

361   in the future. This idea is analogous to recent development in working memory research, which suggests

362   that mnemonic information intended for future use often remains in a latent, activity-silent state (Stokes et

363   al., 2020; Wolff et al., 2015, 2017). Finally, information connectivity between visual and higher-order

364   frontoparietal regions was enhanced by visual impulse during preparation, suggesting improved

365   information flow across the relevant areas allowing enhanced attentional control. This enhanced

366   attentional control may contribute to refined sensory representations of target in early visual cortex,

367   facilitating transitions from a non-sensory to sensory-like format template. Future studies may adopt

368   layer-specific fMRI to infer the direction of this improved information flow (Jia et al., 2023) and explore

17

369 the relationship between long-range connections and the utilization of different versions of target

370 templates.

371 Our study reveals a latent, sensory-like attentional template that extends beyond previous findings

372 of latent working memory representations (Wolff et al., 2015, 2017). Unlike previous studies, which used

373 the "pinging technique" to differentiate active and latent representations of *different* items (e.g., attended

374 vs. unattended memory item), our study demonstrated the active and latent representations of a single

375 item in different formats (i.e., non-sensory vs. sensory-like). Moreover, while the "pinging technique" did

376 not evoke sensory-like neural patterns during memory retention (Wolff et al., 2017), we identified

377 sensory-like format of representations during preparatory attention. A potential neural mechanism for this

378 sensory-like template is via an "activity-silent" mechanism through short-term changes in synaptic

379 weights, originally proposed for working memory storage (Mongillo et al., 2008). However, whether

380 pinging identifies an activity-silent mechanism is currently debated (Barbosa et al., 2021; Schneegans &

381 Bays, 2017). An alternative possibility is that the visual impulse amplified a subtle but active

382 representation of the sensory template during preparation. Distinguishing between these two accounts

383 likely requires neurophysiological measurements, which are beyond the scope of the current study.

384 Nevertheless, our findings provide evidence for dual-format attentional template including a latent,

385 sensory template and offer insights into how the brain encodes information beyond immediate uses to

386 support adaptive behavior.

387 The non-generalizable activity patterns from perception to preparatory attention, in the absence of

388 visual impulse, suggests a default, predominantly non-sensory template during preparation. This finding

389 is largely consistent with electrophysiological studies (Myers et al., 2015; Wen et al., 2019) and our prior

390 fMRI work on preparatory attention to motion directions (Gong et al., 2022), but differs from some

391 previous neuroimaging studies that demonstrated sensory-like templates during preparation (Kok et al.,

392 2014; Peelen & Kastner, 2011; Stokes et al., 2009). One potential account for these discrepancies is that

393 those studies used cue-only trials where the target was expected but not actually presented, contrasting

394 with our task where the target was shown on every trial with temporally separated preparation and

395   stimulus selection periods. This seemingly subtle difference may significantly impact the formats of the

396   neural representations. Because cue-only trials increased the likelihood of target appearance at the

397   subsequent time point, sensory template may be activated due to modulations of temporal expectations

398   (Grubert & Eimer, 2018). This explanation is consistent with theories suggesting differential influences of

399   expectation and attention on neural activity: expectation reflects visual interpretations of stimuli due to

400   sensory uncertainty, whereas attention is guided based on the task relevance of sensory information

401   (Rungratsameetaweemana & Serences, 2019; Summerfield & Egner, 2009, 2016). Our finding of a

402   predominantly non-sensory format may indicate an optimized coding strategy employed by the brain to

403   effectively and robustly represent information for future use. This aligns with the proposed role of

404   attention in modulating sensory representations to encode only currently relevant information at a

405   minimal cost.

406   We acknowledge that our findings cannot pinpoint the exact format of this non-sensory template.

407   Previous behavioral studies on visual search have suggested a candidate mechanism for target-defining

408   features. For instance, when searching for simple features, such as an orientation or a color (Kong et al.,

409   2017; Wolfe et al., 1992), participants tend to use categorical attributes (e.g., steep vs. shallow; left-tilted

410   vs. right-tilted) to enhance search efficiency. The categorical template is particularly advantageous when

411   features are consistent and predictable (Hout et al., 2017). In our task, the angular relations between the

412   target and distractor orientation were defined by categorical attributes (e.g., left-tilted vs. right-tilted) and

413   remained consistent across trials, making a categorical template feasible during preparatory attention.

414   Furthermore, employing a categorical template allows for greater tolerance of stimulus variability, which

415   is also useful in our study where the actual target orientation varied around the reference orientations

416   across trials. Future studies are needed to address the nature of the non-sensory template during

417   preparation as well as task parameters that might modulate them.

418   In summary, the current study suggests that there are two formats of attentional templates each

419   having a distinct functional state: a default, non-sensory format and a latent, sensory-like format. This

420   dual-format representation aligns with theories on the dual-function of attentional template for different

19

421   task goals (Hout & Goldinger, 2015; Yu et al., 2023). The current findings provide a plausible neural

422   implementation for these theories by demonstrating different formats in different functional states. This

423   mechanism likely reflects an optimized coding scheme that effectively balances processing efforts and

424   demands, particularly well-suited for flexible control and transitions from coarse to fine task demands in

425   visually guided behavior.

426

427   Materials and Methods

428   Participants

429   Twenty individuals participated in the No-Ping session (11 females, mean age = 22.9) and twenty

430   individuals participated in the Ping session (14 females, mean age = 23.7). Among them, 14 participants

431   took part in both sessions, while 12 of them took part in only one session. The sample size was

432   comparable to previous studies using similar attention tasks (Baldauf & Desimone, 2014; Gong & Liu,

433   2020a; 2020b; Guo et al., 2012; Jigo et al., 2018; Liu & Hou, 2013). Because our primary interest is the

434   generalization from the perception task to attention task, we used the minimal effect size of decoding

435   accuracy across regions (one-sample t-tests: $d = 0.868$) from our previous study with a similar design

436   (Gong et al., 2022), and used G*Power (Version 3.1) (Faul et al., 2007)to confirm that this sample size is

437   sufficient to detect a cross-task generalization effect with a power greater than 95% (a = 0.05). All

438   participants were right-handed and had a normal or corrected-to-normal vision. Participants provided

439   written informed consent according to the study protocol approved by the Institutional Review Board at

440   Zhejiang University (2020-06-001). They were paid ¥200 (~$27.4) for their participation in each session.

441   Stimuli and Apparatus

442   Stimuli were generated using Psychtoolbox (Brainard, 1997; Kleiner et al., 2007) implemented in

443   MATLAB. The stimuli were presented on an LCD monitor (resolution: 1920 × 1080, refresh rate: 60 Hz)

444   during behavioral training, at a viewing distance of 90 cm in a dark room. During the fMRI scans, stimuli

445   were projected to a screen via a MR-compatible LCD projector (PT-011, Jiexin Technology Co., Ltd.,

446   Shenzhen, China) with the same resolution and refresh rate as the LCD monitor during behavioral training.

447 Participants viewed the screen via an angled mirror attached to the head coil at a viewing distance of 115

448 cm. Angular stimulus size was the same across behavioral and fMRI sessions.

449 The orientation stimuli were square-wave gratings (1.3 cycles per deg, duty cycle: 10%) in a circular

450 aperture (inner radius: 1.5°; outer radius: 6°). The gratings flashed on a gray background at 10 Hz,

451 alternating between black and white. There were two types of stimuli: two overlapping gratings orientated

452 leftward (~135°) and rightward (~45°), or a single grating with one of the two orientations (~135° or

453 ~45°). Here, we refer to the 45° and 135° orientations as the reference orientations. The impulse stimulus

454 was a high-contrast, white (at the maximum projector output level) circular disk that covered the same

455 area as the orientation stimulus (radius: 6°).

456 **Experimental Procedures and Tasks**

457 Each participant completed at least two fMRI sessions on different days. One session was used for

458 defining ROIs (see Definition of Regions of Interests), while the remaining sessions were used for the

459 main experiment (see Attention task and Perception task). Before the scanning sessions, participants were

460 trained to familiarize with the tasks in a separate behavioral session. The procedures and tasks were

461 similar to our previous work (Jigo et al., 2018; Gong et al., 2022).

462 *Attention task.* We used a cueing paradigm (Fig 1A). Each trial began with a color cue (red or

463 green) for 0.5 s to indicate the reference orientation of the upcoming target (leftward vs. rightward

464 orientation). In the No-Ping session, the cue was followed by a blank display during the preparation

465 period; in the Ping session, a task-irrelevant, high-luminance visual impulse ("ping", 0.1 s) occurred at

466 either 0.5 s (for short delays of 1.5 s and 3.5 s) or 2.5 s (for long delays of 5.5 s and 7.5 s) after the onset

467 of the cue display. The orders of these sessions were counterbalanced across participants who completed

468 both. Following the preparatory period, two superimposed gratings were then shown for 1 s. The target

469 grating was shown with a small angular offset with respect to the cued reference orientation, whereas the

470 distractor grating was shown in the uncued reference orientation (e.g., if rightward orientation was cued,

471 the rightward grating was shown in 45°± d and the leftward grating was shown in 135°). Note that the

472 angular offset was determined individually based on the threshold obtained during the training session (at

21

473   least 3 blocks, 30 trials/block), using a staircase procedure (Best Parameter Estimation by Sequential

474   Testing, Best PEST), as implemented in the Palamedes Toolbox (Prins & Kingdom, 2009). Participants

475   used a keypad to report whether the attended orientation was more leftward or rightward relative to the

476   reference orientation. Each trial was separated by an inter-trial interval of 3 s to 7 s (2 s per step). Trial-

477   by-trial feedback ("correct" or "incorrect") was provided in the training session but not during scanning.

478   Instead, the percentage of correct responses was provided at the end of each run in the scanning session to

479   avoid the impact of trial-level feedback on neural activity.

480   To prevent the cue-related sensory difference from contributing to neural activity, we reversed the

481   mapping between colors and orientations halfway through the experiment (e.g., red indicated "attend

482   leftward orientation" and green indicated "attend rightward orientation" in the first half of the runs, and

483   vice versa for the second half of the runs), with the order counterbalanced across subjects. The mapping

484   of colors and orientations was reversed only once in the middle of the experiment to prevent

485   misremembering of the color-orientation associations. To reduce temporal expectancy over a fixed period,

486   the preparatory period (i.e., cue-to-stimulus interval) varied from 1.5 s to 7.5 s with different probabilities

487   (10% for 1.5 s or 3.5 s each, 40% for 5.5 s or 7.5 s each). The long-delay trials (5.5 s or 7.5 s) were

488   selected for subsequent analyses, as they allow the separation of the preparatory activity from the grating-

489   evoked response during fMRI scanning. The short-delay trials were included to encourage a sustained

490   maintenance of attention throughout the entire preparation period.

491   *Perception task.* On each trial of the perception task (Fig 1B), a single grating was shown for 1 s,

492   followed by an inter-trial interval between 3 s to 7 s. To equate the sensory inputs between attention and

493   perception tasks, the orientation was shifted away from the reference orientation by the same angular

494   offset as that used in the attention task with each individual participant's own threshold. Participants

495   performed the same orientation discrimination task by comparing the single orientation to the reference

496   orientation. We provided the percentage of correct responses at the end of each run as feedback.

497   **Eye Tracking and Analysis**

22

498   To evaluate the stability of visual fixation, we used Eyelink Portable Duo system (SR Research, Ontario,

499   Canada) to monitor each observer's eye position during the training session at a sampling rate of 500 Hz.

500   One participant's data was not used due to the unstable recording of the eye. The data were then analyzed

501   using custom Matlab code.

502   To examine whether participants adopted a space-based strategy during the preparatory period in

503   the attention task, such as directing their gaze leftward in attend leftward trials, and vice versa for the

504   attend rightward trials, we analyzed the participants' eye positions recorded during the training session.

505   Horizontal and vertical eye positions were analyzed separately. Paired t-tests were performed to compare

506   horizontal and vertical eye positions between two attention conditions. A two-way mixed ANOVA (2

507   sessions $\times$ 2 attended orientations) was applied to the horizontal and vertical positions, respectively.

508   **fMRI Data Acquisition**

509   Imaging was performed on a Siemens 3T scanner (MAGNETOM Prisma, Siemens Healthcare, Erlangen,

510   Germany) using a 20-channel coil at Zhejiang University (Hangzhou, China). For each participant, we

511   acquired high-resolution T1-weighted anatomical images (field of view, 240 $\times$ 240 mm, 208 sagittal

512   slices; 0.9 mm^3 resolution), T2*-weighted echo-planar functional images consisting of 46 slices (TR, 2 s;

513   TE, 34 ms; flip angle, 50°; matrix size, 80 $\times$ 80; in-plane resolution, 3 $\times$ 3 mm; slice thickness, 3 mm,

514   interleaved, no gap) and a 2D T1-weighted anatomical image (0.8 $\times$ 0.8 $\times$ 3 mm) for aligning functional

515   data to high-resolution anatomical data.

516   **fMRI Data Preprocessing**

517   Data analyses were performed using mrTools (Gardner et al., 2018) and custom code in Matlab. For each

518   run, functional data were preprocessed with head motion correction, linear detrending and temporal high

519   pass filtering at 0.01 Hz. Data were converted to percentage signal change by dividing the time course of

520   each voxel by its mean signals in each run. We concatenated the 6 runs of the attention task and the 3 runs

521   of the perception task separately for further analysis. One of the attention runs in one subject was

522   excluded due to low accurate performance (<50%).

Definition of Regions of Interests (ROI)

Visual and parietal ROIs. Following previous work (Jigo et al., 2018; Gong et al., 2020b; Gong et al., 2022), for each observer, we ran a separate retinotopic mapping session to obtain ROIs in occipital and parietal areas. Observers viewed four runs of rotating wedges (i.e., clockwise and counterclockwise) and two runs of rings (i.e., expanding and contracting) to map the polar angle and radial components, respectively (DeYoe et al., 1996; Engel, 1997; Sereno et al., 1995). Borders between areas were defined as the phase reversals in a polar angle map of the visual field. Phase maps were visualized on computationally flattened representations of the cortical surface, which were generated from the high-resolution anatomical image using FreeSurfer (http://surfer.nmr.mgh.harvard.edu) and custom Matlab code.

To help identify the topographic areas in parietal areas, we ran 2 runs of memory-guided saccade task modeled after previous studies (Konen & Kastner, 2008; Schluppeck et al., 2006; Sereno et al., 2001). Observers fixated at the screen center while a peripheral (~10° radius) target dot was flashed for 500 ms. The flashed target was quickly masked by a ring of 100 distractor dots randomly positioned within an annulus (8.5° – 10.5°). The mask remained on screen for 3 s, after which participants were instructed to make a saccade to the memorized target position, then immediately saccade back to the central fixation. The position of the peripheral target shifted around the annulus from trial to trial in either a clockwise or counterclockwise order. Data from the memory-guided saccade task were analyzed using the same phase encoding method as the wedge and ring data. Therefore, the following regions of interest (ROIs) in each hemisphere were identified after the completion of this session: V1, V2, V3, V3A/B, V4, V7/IPS0, IPS1 to IPS4.

Frontal ROIs. Following previous work (Jigo et al., 2018; Gong et al., 2020b; Gong et al., 2022), we used a deconvolution approach by fitting each voxel's time series from the attention task with a general linear model (GLM) to determine the event-related activations in the brain (see Supplementary Materials: Deconvolution). For each voxel, we computed the goodness of fit measure (r^2 value), which indicates the amount of variance explained by the deconvolution model (Gardner et al., 2018). The r^2

24

549 value represents the degree to which the voxel's time series is correlated with the task events, regardless

550 of any differential responses among conditions. Based on the task-related activation (as indexed by r^2

551 value) and anatomical criteria, we defined two frontal areas in each hemisphere that were active during

552 the attention task: one is located superior to the precentral sulcus and near the superior frontal sulcus (FEF)

553 and the other is located towards the inferior precentral sulcus, close to the junction with the inferior

554 frontal sulcus (IFJ).

555 **Groups of Region.** To characterize the patterns of neural response across cortical hierarchy and

556 streamline data presentation, we grouped results from the nine areas into four groups based on functional

557 and anatomical considerations: primary visual cortex (V1); extrastriate visual cortex (EVC), consisting of

558 V2, V3, V3ab and V4; IPS, consisting of IPS0 to IPS4; prefrontal cortex (PFC), consisting of FEF and

559 IFJ. Individual areas within each group exhibited qualitatively similar results.

560 **Multivoxel Pattern Analysis (MVPA)**

561 **Decoding of attended orientation.** To test if multivariate patterns of activity represent

562 information of the attended orientation, separate MVPA analyses were applied on the activity patterns for

563 the preparation and stimulus selection periods. Following previous work (Jigo et al., 2018; Gong et al.,

564 2020b; Gong et al., 2022), for this analysis, we extracted fMRI signals from raw time series in the long

565 delay trials with correct behavioral responses (~72 trials per attention condition); short delay trials were

566 excluded as they could not provide enough data points to measure preparatory activity. We then obtained

567 averaged BOLD response in a 2 s window for each voxel and each trial in a given ROI, separately for

568 preparatory activity (4 to 6 s after the onset of the cue) and stimulus-evoked activity (4 to 6 s after the

569 onset of the gratings). The response amplitudes across two attention conditions in each ROI were further

570 z-normalized, separately for the preparation and stimulus-related activity. These normalized single-trial

571 BOLD responses were used for the MVPA. We trained a classifier using the Fisher linear discriminant

572 (FLD) analysis to discriminate between two attended orientations (leftward vs. rightward) and tested its

573 performance with a leave-one-run-out cross-validation scheme. This process was repeated until each run

574 was tested once and the decoding accuracy (i.e., the proportion of correctly classified trials) was averaged

25

575 across the cross-validation folds. The statistical significance of decoding accuracy was evaluated by

576 comparing it to the chance level obtained from a permutation test (see Permutation test). To assess if the

577 decoding accuracy differed between No-Ping and Ping experiments, we performed two-way mixed

578 ANOVAs (2 sessions × 4 regions) on the decoding accuracy.

579   **Cross-task generalization from the perception task to attention task.** Following previous

580 work (Jigo et al., 2018; Gong et al., 2022), to test whether the neural patterns in the preparatory and

581 stimulus selection periods from the attention task reflected sensory processing of isolated features, we

582 trained an FLD classifier using the normalized BOLD responses from the perception task (4 to 6 s after

583 the trial onset) to discriminate leftward vs. rightward orientation. Then, we tested this classifier on the

584 normalized response from the independent runs of attention task to discriminate between attend leftward

585 vs. attend rightward orientations, separately for preparation and stimulus selection periods. The

586 significance of decoding accuracy was compared to the chance level obtained from a permutation test (see

587 Permutation test). To assess if the generalization performance differed between No-Ping and Ping

588 sessions, we performed two-way mixed ANOVAs (2 sessions × 4 regions) on the decoding accuracy.

589   **Neural distance between attended and perceived orientations.** The decoding accuracy from

590 the cross-task generalization test reflects a discretized readout of the pattern similarity between different

591 conditions. However, employing continuous similarity measures, such as Mahalanobis distance

592 (Mahalanobis, 1936), could be more reliable compared to decoding accuracy (Walther et al., 2016).

593 Therefore, we calculated the Mahalanobis distance to quantify the pattern similarity between two attended

594 orientations and two perceived orientations. For each participant and each ROI, we have M points (i.e., M

595 trials for each attended orientation) in the N-dimensional space (N = 100, number of voxels). For each

596 data point in the attended orientation condition, we computed its distance to each of the orientation

597 distributions (from the perception task). Averaged distance values were then calculated for each

598 combination of attended orientation and perceived orientation pairs. A sensory-like hypothesis would

599 predict smaller distance between the distribution of the attended orientation (e.g., attend leftward) and the

600 distribution of corresponding orientation (e.g., perceive leftward) compared to the alternative orientation

601   (e.g., perceive rightward). Two-way repeated-measures ANOVA (2 attended orientations × 2 perceived

602   orientations) was applied on the Mahalanobis distance, separately for each region and each session.

603   **Neural-behavioral relationships.** We tested if the representation format during preparatory

604   attention was associated with subsequent behavior. For each trial, we calculated the Mahalanobis distance

605   between the attention conditions (attend leftward and attend rightward) and the perceived orientations

606   (leftward and rightward orientation). We estimated the attentional modulation index (AMI) based on these

607   distance values. This index measures how much attention modulated the pattern similarity for the Same

608   orientation condition (e.g., attend and perceive the same orientation) relative to the Different orientation

609   condition (e.g., attend and perceive different orientations). The index was calculated as follows: AMI =

610    $(D_{different} - D_{same})/(D_{different} + D_{same})$, where D_{same} and $D_{different}$ are the measured distance (D) in the Same and

611   Different orientation condition, respectively. Next, we sorted the single-trial AMI values in descending

612   order and selected top-ranked 25% trials and bottom-ranked 25% trials to represent "strong modulation"

613   and "weak modulation" trials, respectively. We then extracted behavioral responses on these selected

614   trials and calculated the reaction time and accuracy for each trial type. Paired t-tests were used to compare

615   between "strong modulation" and "weak modulation" trials in each session.

616   **Informational connectivity analysis.** We used Informational Connectivity (IC) to examine

617   shared changes in pattern discriminability over time, a method that allows inference based on multivoxel

618   pattern information rather than overall BOLD response (Jia et al., 2020; Ng et al., 2021). To track the

619   flow of multivariate information across time (i.e., across trials), we measured the fluctuations (covariance)

620   in pattern-based discriminability by calculating the Mahalanobis distance of each trial to the two attended

621   orientations, using a leave-one-run-out cross-validation scheme. For each ROI, we calculated the

622   Mahalanobis distance between the pattern of activity for each attention trial from one left-out run and the

623   distribution of each attended orientation of the remaining runs. To quantify the degree of attentional

624   modulation, we calculated the AMI using the same formula as mentioned above, where D_{same} and $D_{different}$

625   are the measured distance (D) in the Same and Different condition. This index measures how much the

626   pattern similarity increased for the same attention condition (e.g., attend leftward to attend leftward)

27

627   relative to the different attention condition (e.g., attend leftward to attend rightward). A positive AMI

628   indicates relative proximity to the same attention condition, whereas a negative AMI indicates relative

629   proximity to the different attention condition. A time course of AMI values was generated across runs and

630   pairwise correlated between ROIs using Pearson correlation analysis and Fisher z-transformed.

631   Independent t-tests were used to compare the connectivity between No-Ping and Ping sessions.

632   **Permutation test to evaluate classifier performance**

633   Following previous work (Jigo et al., 2018; Gong et al., 2022), for each brain area, we evaluated the

634   statistical significance of the observed decoding accuracy using a permutation test scheme. We first

635   shuffled the trial labels in the training data and trained the same FLD classifier on the shuffled data. We

636   then tested the classifier on the (unshuffled) test data to obtain decoding accuracy. For each ROI and each

637   participant, we repeated this procedure 1000 times to compute a null distribution of decoding accuracy.

638   To compute the group-level significance, we averaged the 20 null distributions to obtain a single null

639   distribution of 1000 values for each ROI. To determine if the observed decoding accuracy significantly

640   exceeds the chance level, we compared the observed value to the 95 percentiles of this group-level

641   distribution (corresponding to $p = 0.05$). Note that these ROIs were pre-defined with strong priors as their

642   activation in attention tasks has been consistently reported in the literature. Nevertheless, for those

643   analyses where multiple comparisons were performed across regions, we applied a Bonferroni-correction

644   to adjust the p-values.

645   **Bayesian analysis**

646   To evaluate the strength of evidence for the null hypothesis, we conducted Bayesian analyses

647   (Wagenmakers, 2007) using standard priors as implemented in JASP Version 0.17.1 (JASP Team, 2023).

648   We performed Bayesian t-tests and computed Bayes factor (BF_{01}) to compare between two attention

649   conditions (attend leftward vs. attend rightward). Additionally, we used Bayesian repeated-measures

650   ANOVA and computed the exclusion Bayes factors (BF_{excl}) to assess the evidence for excluding specific

651   effects across all models. A Bayes factor (BF) greater than 1 provides support for the null hypothesis.

28

652   Specifically, a BF between 1 and 3 indicates weak evidence, a BF between 3 and 10 indicates moderate

653   evidence, and a BF greater than 10 indicates strong evidence (Van Doorn et al., 2021).

654   **Approach to handle partially overlapped samples**

655   Our study used partially overlapping samples, with 14 out of 20 participants completing both No-Ping

656   and Ping sessions, while the remainder completed one of the two sessions. The most important analyses

657   entailed assessing whether decoding accuracy was above chance, for which we used the permutation-

658   based method (see above) within each session. Thus, these analyses were unaffected by the partially

659   overlapping samples. In a few analyses where we compared across sessions, we used statistical tests

660   treating "session" as a between-subject factor. We believe this is a reasonable approach, as a between-

661   subject test is more conservative than a within-subject test, such that any significant effect emerged

662   should be a genuine effect. To be certain, we also conducted additional analyses with "session" as a

663   within-subject factor on the subset of data from the 14 participants who completed both sessions in a

664   counterbalanced order. The results were highly similar to those reported in the main text.

References

Baldauf, D., & Desimone, R. (2014). Neural mechanisms of object-based attention. *Science*, *344*(6182), 424–427. https://doi.org/10.1126/science.1247003

Barbosa, J., Lozano-Soldevilla, D., & Compte, A. (2021). Pinging the brain with visual impulses reveals electrically active, not activity-silent, working memories. *PLOS Biology*, *19*(10), e3001436. https://doi.org/10.1371/journal.pbio.3001436

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Bressler, S. L., Tang, W., Sylvester, C. M., Shulman, G. L., & Corbetta, M. (2008). Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *The Journal of Neuroscience*, *28*(40), 10056–10061. https://doi.org/10.1523/JNEUROSCI.1776-08.2008

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., Miller, D., & Neitz, J. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences*, *93*(6), 2382–2386. https://doi.org/10.1073/pnas.93.6.2382

Duncan, D. H., van Moorselaar, D., & Theeuwes, J. (2023). Pinging the brain to reveal the hidden attentional priority map using encephalography. *Nature Communications*, *14*(1), 4749. https://doi.org/10.1038/s41467-023-40405-8

Engel, S. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, *7*(2), 181–192. https://doi.org/10.1093/cercor/7.2.181

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

30

691   Gardner, J. L., Merriam, E. P., Schluppeck, D., Besle, J., & Heeger, D. J. (2018). mrTools: Analysis and

692   visualization package for functional magnetic resonance imaging data (Version 4.7) [Computer

693   software]. Zenodo. https://doi.org/10.5281/ZENODO.1299483

694   Gong, M., Chen, Y., & Liu, T. (2022). Preparatory attention to visual features primarily relies on non-

695   sensory representation. *Scientific Reports*, *12*(1), 21726. https://doi.org/10.1038/s41598-022-26104-

696   2

697   Gong, M., & Liu, T. (2020a). Biased neural representation of feature-based attention in the human

698   frontoparietal network. *Journal of Neuroscience*, *40*(43), 8386-8395.

699   Gong, M., & Liu, T. (2020b). Continuous and discrete representations of feature-based attentional priority

700   in human frontoparietal network. *Cognitive Neuroscience*, *11*(1–2), 47–59.

701   https://doi.org/10.1080/17588928.2019.1601074

702   Grubert, A., & Eimer, M. (2018). The time course of target template activation processes during

703   preparation for visual search. *The Journal of Neuroscience*, *38*(44), 9527–9538.

704   https://doi.org/10.1523/JNEUROSCI.0409-18.2018

705   Guo, F., Preston, T. J., Das, K., Giesbrecht, B., & Eckstein, M. P. (2012). Feature-independent neural

706   coding of target detection during search of natural scenes. *Journal of Neuroscience*, *32*(28), 9499–

707   9510. https://doi.org/10.1523/JNEUROSCI.5876-11.2012

708   Hamblin-Frohman, Z., & Becker, S. I. (2021). The attentional template in high and low similarity search:

709   Optimal tuning or tuning to relations? *Cognition*, *212*, 104732.

710   https://doi.org/10.1016/j.cognition.2021.104732

711   Hout, M. C., & Goldinger, S. D. (2015). Target templates: The precision of mental representations affects

712   attentional guidance and decision-making in visual search. *Attention, Perception, & Psychophysics*,

713   *77*(1), 128–149. https://doi.org/10.3758/s13414-014-0764-6

714   Hout, M. C., Robbins, A., Godwin, H. J., Fitzsimmons, G., & Scarince, C. (2017). Categorical templates

715   are more useful when features are consistent: Evidence from eye movements during search for

31

716 societally important vehicles. *Attention, Perception, & Psychophysics*, *79*(6), 1578–1592.

717 https://doi.org/10.3758/s13414-017-1354-1

718 JASP Team. (2023). JASP (Version 0.17.1) [Computer software].

719 Jia, K., Goebel, R., & Kourtzi, Z. (2023). Ultra-high field imaging of human visual cognition. *Annual*

720 *Review of Vision Science*, *9*(1), 479–500. https://doi.org/10.1146/annurev-vision-111022-123830

721 Jia, K., Zamboni, E., Kemper, V., Rua, C., Goncalves, N. R., Ng, A. K. T., Rodgers, C. T., Williams, G.,

722 Goebel, R., & Kourtzi, Z. (2020). Recurrent processing drives perceptual plasticity. *Current Biology*,

723 *30*(21), 4177-4187.e4. https://doi.org/10.1016/j.cub.2020.08.016

724 Jigo, M., Gong, M., & Liu, T. (2018). Neural determinants of task performance during feature-based

725 attention in human cortex. *eNeuro*, *5*(1), 1–15. https://doi.org/10.1523/ENEURO.0375-17.2018

726 Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in

727 human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*(4),

728 751–761. https://doi.org/10.1016/S0896-6273(00)80734-5

729 Kerzel, D. (2019). The precision of attentional selection is far worse than the precision of the underlying

730 memory representation. *Cognition*, *186*, 20–31. https://doi.org/10.1016/j.cognition.2019.02.001

731 Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3?

732 Kok, P., Failing, M. F., & De Lange, F. P. (2014). Prior expectations evoke stimulus templates in the

733 primary visual cortex. *Journal of Cognitive Neuroscience*, *26*(7), 1546–1554.

734 https://doi.org/10.1162/jocn_a_00562

735 Konen, C. S., & Kastner, S. (2008). Representation of eye movements and stimulus motion in

736 topographically organized areas of human posterior parietal cortex. *Journal of Neuroscience*, *28*(33),

737 8361–8375. https://doi.org/10.1523/JNEUROSCI.1930-08.2008

738 Kong, G., Alais, D., & Van Der Burg, E. (2017). Orientation categories used in guidance of attention in

739 visual search can differ in strength. *Attention, Perception, & Psychophysics*, *79*(8), 2246–2256.

740 https://doi.org/10.3758/s13414-017-1387-5

741   Lewis-Peacock, J. A., Drysdale, A. T., & Postle, B. R. (2015). Neural evidence for the flexible control of

742   mental representations. *Cerebral Cortex*, *25*(10), 3303–3313. https://doi.org/10.1093/cercor/bhu130

743   Liu, T., & Hou, Y. (2013). A hierarchy of attentional priority signals in human frontoparietal cortex.

744   *Journal of Neuroscience*, *33*(42), 16606–16616. https://doi.org/10.1523/JNEUROSCI.1780-13.2013

745   Mahalanobis, P. C. (1936). On generalised distance in statistics. *National Institute of Science — 3India,*

746   *Proceedings*, *2*, 49–55.

747   Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in

748   real-world scenes: Evidence from eye movements. *Journal of Vision*, *9*(11), 8–8.

749   https://doi.org/10.1167/9.11.8

750   Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, *319*(5869),

751   1543–1546. https://doi.org/10.1126/science.1150769

752   Myers, N. E., Walther, L., Wallis, G., Stokes, M. G., & Nobre, A. C. (2015). Temporal dynamics of

753   attention during encoding versus maintenance of working memory: complementary views from

754   event-related potentials and alpha-band oscillations. *Journal of Cognitive Neuroscience*, *27*(3), 492–

755   508. https://doi.org/10.1162/jocn_a_00727

756   Ng, A. K. T., Jia, K., Goncalves, N. R., Zamboni, E., Kemper, V. G., Goebel, R., Welchman, A. E., &

757   Kourtzi, Z. (2021). Ultra-High-Field neuroimaging reveals fine-scale processing for 3D perception.

758   *The Journal of Neuroscience*, *41*(40), 8362–8374. https://doi.org/10.1523/JNEUROSCI.0065-

759   21.2021

760   Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human

761   occipitotemporal cortex. *Proceedings of the National Academy of Sciences of the United States of*

762   *America*, *108*(29), 12125–12130. https://doi.org/10.1073/pnas.1101042108

763   Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data.

764   *Palamedes: Matlab Routines for Analyzing Psychophysical Data.*

765   Rosenberg, M. D., Scheinost, D., Greene, A. S., Avery, E. W., Kwon, Y. H., Finn, E. S., Ramani, R., Qiu,

766   M., Constable, R. T., & Chun, M. M. (2020). Functional connectivity predicts changes in attention

767   observed across minutes, days, and months. *Proceedings of the National Academy of Sciences*,

768   *117*(7), 3797–3807. https://doi.org/10.1073/pnas.1912226117

769   Rungratsameetaweemana, N., & Serences, J. T. (2019). Dissociating the impact of attention and

770   expectation on early sensory processing. *Current Opinion in Psychology*, *29*, 181–186.

771   https://doi.org/10.1016/j.copsyc.2019.03.014

772   Schluppeck, D., Curtis, C. E., Glimcher, P. W., & Heeger, D. J. (2006). Sustained activity in topographic

773   areas of human posterior parietal cortex during memory-guided saccades. *The Journal of*

774   *Neuroscience*, *26*(19), 5098–5108. https://doi.org/10.1523/JNEUROSCI.5330-05.2006

775   Schneegans, S., & Bays, P. M. (2017). Restoration of fMRI decodability does not imply latent working

776   memory states. *Journal of Cognitive Neuroscience*, *29*(12), 1977–1994.

777   https://doi.org/10.1162/jocn_a_01180

778   Scolari, M., Byers, A., & Serences, J. T. (2012). Optimal deployment of attentional gain during fine

779   discriminations. *The Journal of Neuroscience*, *32*(22), 7723–7733.

780   https://doi.org/10.1523/JNEUROSCI.5558-11.2012

781   Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., &

782   Tootell, R. B. H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic

783   resonance imaging. *Science*, *268*(5212), 889–893. https://doi.org/10.1126/science.7754376

784   Sereno, M. I., Pitzalis, S., & Martinez, A. (2001). Mapping of contralateral space in retinotopic

785   coordinates by a parietal cortical area in humans. *Science*, *294*(5545), 1350–1354.

786   https://doi.org/10.1126/science.1063695

787   Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: A dynamic coding

788   framework. *Trends in Cognitive Sciences*, *19*(7), 394–405. https://doi.org/10.1016/j.tics.2015.05.004

789   Stokes, M. G., Muhle-Karbe, P. S., & Myers, N. E. (2020). Theoretical distinction between functional

790   states in working memory and their corresponding neural states. *Visual Cognition*, *28*(5–8), 420–432.

791   https://doi.org/10.1080/13506285.2020.1825141

792   Stokes, M., Thompson, R., Nobre, A. C., & Duncan, J. (2009). Shape-specific preparatory activity

793   mediates attention to targets in human visual cortex. *Proceedings of the National Academy of*

794   *Sciences*, *106*(46), 19569–19574. https://doi.org/10.1073/pnas.0905306106

795   Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: Neural and

796   computational mechanisms. *Nature Reviews Neuroscience*, *15*(11), 745–756.

797   https://doi.org/10.1038/nrn3838

798   Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive*

799   *Sciences*, *13*(9), 403–409. https://doi.org/10.1016/j.tics.2009.06.003

800   Summerfield, C., & Egner, T. (2016). Feature-based attention and feature-based expectation. *Trends in*

801   *Cognitive Sciences*, *20*(6), 401–404. https://doi.org/10.1016/j.tics.2016.03.008

802   Van Doorn, J., Van Den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J.,

803   Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Gupta, A. R.

804   K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E.-J. (2021). The JASP guidelines

805   for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*(3), 813–826.

806   https://doi.org/10.3758/s13423-020-01798-5

807   Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems ofp values. *Psychonomic*

808   *Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

809   Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of

810   dissimilarity measures for multi-voxel pattern analysis. *NeuroImage*, *137*, 188–200.

811   https://doi.org/10.1016/j.neuroimage.2015.12.012

812   Wen, T., Duncan, J., & Mitchell, D. J. (2019). The time-course of component processes of selective

813   attention. *NeuroImage*, *199*, 396–407. https://doi.org/10.1016/j.neuroimage.2019.05.067

814   Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin &*

815   *Review*, *28*(4), 1060–1092. https://doi.org/10.3758/s13423-020-01859-9

816   Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., & O'Connell, K. M. (1992). The role of categorization

817   in visual search for orientation. *Journal of Experimental Psychology: Human Perception and*

818   *Performance*, *18*(1), 34–49. https://doi.org/10.1037/0096-1523.18.1.34

819   Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working

820   memory using electroencephalography. *Frontiers in Systems Neuroscience*, *9*.

821   https://doi.org/10.3389/fnsys.2015.00123

822   Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying

823   working-memory-guided behavior. *Nature Neuroscience*, *20*(6), 864–871.

824   https://doi.org/10.1038/nn.4546

825   Yu, X., Hanks, T. D., & Geng, J. J. (2022). Attentional guidance and match decisions rely on different

826   template information during visual search. *Psychological Science*, *33*(1), 105–120.

827   https://doi.org/10.1177/09567976211032225

828   Yu, X., Zhou, Z., Becker, S. I., Boettcher, S. E. P., & Geng, J. J. (2023). Good-enough attentional

829   guidance. *Trends in Cognitive Sciences*, *27*(4), 391–403. https://doi.org/10.1016/j.tics.2023.01.007

830   Zhang, H., & Luo, H. (2023). Feature-specific reactivations of past information shift current neural

831   encoding thereby mediating serial bias behaviors. *PLOS Biology*, *21*(3), e3002056.

832   https://doi.org/10.1371/journal.pbio.3002056

833

Acknowledgements

Author contributions

Original idea by K.J. and M.G. Experiment programmed by Y.C. Data collected by Y.C. Analysis programmed by Y.C., K.J. and M.G. Manuscript drafted by Y.C., T.L., and M.G. Manuscript edited and revised by T.L., J. T., K.J., and M.G.

Data Availability

All data, analyses, and task codes have been made publicly available via the Open Science Framework at

https://osf.io/ghaxv/

Competing interests

The authors declare no competing interests.