

1 Running Head: Attention and Working Memory

2 Reorienting Spatial Attention within Visual Working Memory

3

4 Sizhu Han^{1,} , Yixuan Ku^{1,2,*}

5

6 1. The Shanghai Key Lab of Brain Functional Genomics, Shanghai
7 Changning-ECNU Mental Health Center, School of Psychology and Cognitive
8 Science, East China Normal University, Shanghai, China

9 2. NYU-ECNU Institute of Brain and Cognitive Science, NYU Shanghai and
10 Collaborative Innovation Center for Brain Science, Shanghai, China

11

12 * correspondence: Yixuan Ku, kuyixuan@gmail.com

13

14 **Abstract**

15 Attention and working memory (WM) are intertwined core cognitive processes.
 16 Through four experiments with 133 participants, we dissociated the impact of
 17 two types of covert spatial attention, endogenous vs. exogenous, on visual WM.
 18 Behavioral results consistently indicated that exogenous attentional cues were
 19 more advantageous than endogenous ones in enhancing the precision of visual
 20 WM under load-2, while they equalized under load-4. In addition, physiological
 21 and neural data explained the mechanisms. Converging evidence from
 22 eye-tracking, electroencephalography, and magnetoencephalography
 23 suggested that fast attentional processing induced by exogenous cues lead to
 24 early top-down information from the dorsal lateral prefrontal cortex (DLPFC) to
 25 sensory cortices. The differential frontal activities were further correlated with
 26 the behavioral distinctions between exogenous and endogenous cues, and
 27 transcranial magnetic stimulation over DLPFC at the same time period
 28 abolished the exogenous advantage. Taken together, traditionally considered
 29 bottom-up attentional processing induced by exogenous cues rapidly engages
 30 top-down signals from the frontal cortex, which leads to stronger behavioral
 31 benefits compared with the benefits produced by endogenous cues under the
 32 low load condition.

33

34 **Keywords:** Visual working memory; spatial attention; exogenous; endogenous;
 35 top-down control

36

37

38 INTRODUCTION

39 Working memory (WM) is a critical ability to store and manipulate sensory
40 information when it is no longer accessible in the environment (Baddeley, 1986).
41 Visual working memory (VWM) capacity is severely limited (Cowan, 2001;
42 Miller, 1956), but lies in the center of cognition and is linked to the general
43 intelligence (Engle, Kane, & W Tuholski, 1999). VWM representations,
44 however, are flexible (Van den Berg, Awh, & Ma, 2014) and could be influenced
45 by retro-cues (Griffin & Nobre, 2003; Souza & Oberauer, 2016), indicating
46 interactions between attention and VWM.

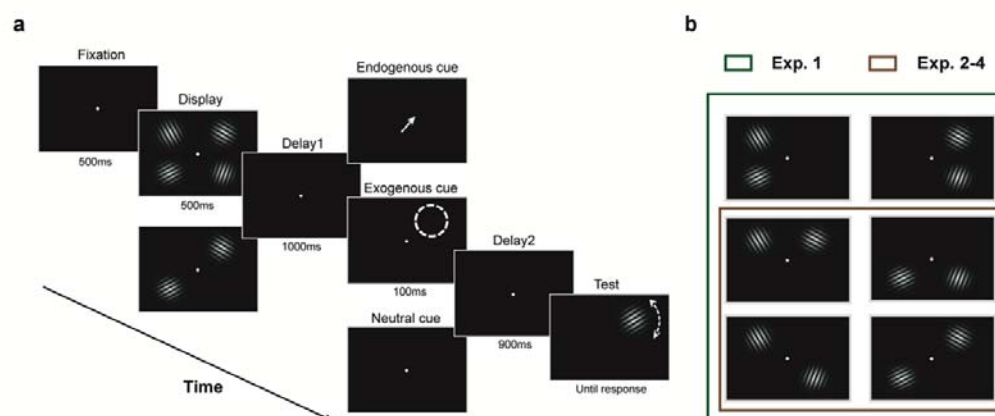
47 As complex as memory, attention also has many different sub-types from
48 distinct standards of classification (Cowan, 2017; Carrasco, 2011; Fan,
49 McCandliss, Fossella, Flombaum, & Posner, 2005; Posner, 1980). Since
50 William James, attention has been divided into two types: one that is reflexive
51 and involuntary, and another that is active and voluntary (James, Burkhardt,
52 Bowers, & Skrupskelis, 1890; Johnston & Dark, 1986). These two kinds of
53 attention correspond to bottom-up and top-down attention from neural views
54 (Buschman & Miller, 2007; Hahn, Ross, & Stein, 2006; Theeuwes, 2010).
55 Covert attention, as opposed to overt attention, is the attentional deployment
56 towards peripheral locations while the eyes gaze at the center. It can be
57 directed by exogenous or endogenous cues, corresponding to the above
58 reflexive/involuntary/bottom-up and active/voluntary/top-down ways
59 respectively.

60 In the perception domain when the attentional cues are given prior to the
61 sensory stimuli, a plethora of studies indicated that the exogenous attention
62 processes rise (~100ms) and decay quickly (Liu, Stevens, & Carrasco, 2007;
63 Remington, Johnston, & Yantis, 1992), while the endogenous attention
64 processes deploy slower (~300ms) but remains sustained (Busse, Katzner, &
65 Treue, 2008). However, the two attentional effects over perception are
66 normally equivalent given enough time for attention to deploy (See review in
67 Carrasco, 2011), although there exist part of evidence claiming their behavioral
68 difference when sensory signal-to-noise ratio is high (Lu & Doshier, 1998, 2000;
69 Lu, Lesmes, & Doshier, 2002). Despite of their similar behavioral benefit, the
70 two types of cues have already been suggested to induce neural processes

71 that involve overlapped but distinctive neural networks (Buschman & Miller,
72 2007) (Chica, Bartolomeo, & Lupiáñez, 2013; Carrasco, 2011; Corbetta &
73 Shulman, 2002), supporting the dissociated neural view of bottom-up vs.
74 top-down.

75 Meanwhile, in the memory domain when the attentional cues are given
76 during the delay period after the sensory stimuli disappear, many studies have
77 suggested that the two attentional effects over memory representations can
78 enhance behavioral performance (e.g., Astle, Summerfield, Griffin, & Nobre,
79 2012; Brady & Hampton, 2018; Gözenman, Tanoue, Metoyer, & Berryhill, 2014;
80 Gunseli, van Moorselaar, Meeter, & Olivers, 2015; Lepsien & Nobre, 2006;
81 Murray, Nobre, Clark, Cravo, & Stokes, 2013; Williams, Hong, Kang, Carlisle, &
82 Woodman, 2013; Griffin & Nobre, 2003; Matsukura, Cosman, Roper, Vatterott,
83 & Vecera, 2014; Pertzov et al., 2013; Tanoue & Berryhill, 2012). However, none
84 of them have found significant difference between them. There might be
85 several reasons for such null effects. First, insensitive binary report with a
86 change detection task (Shimi, Nobre, Astle, & Scerif, 2014) may not catch their
87 subtle difference. Second, relatively high task difficulty (load 4) with a
88 free-recall task (Pertzov et al., 2013) is similar to the condition of low
89 signal-to-noise ratio in the perception domain. Increasing the memory
90 signal-to-noise ratio might reveal the distinctions in the memory domain, given
91 several findings of the differences in the perceptual domain when the
92 sensory-to-noise ratio is high (Lu & Doshier, 1998, 2000; Lu, Lesmes, & Doshier,
93 2002). Third, the effects of endogenous and exogenous cues in the memory
94 domain are truly similar after excluding the above two reasons. Nevertheless,
95 even if the endogenous and exogenous cues produce similar behavioral results
96 in the memory domain, they may induce different neural processes and involve
97 different neural networks, seeing that neural networks induced by them in the
98 perceptual domain are distinctive (Buschman & Miller, 2000; Chica,
99 Bartolomeo, & Lupiáñez, 2013; Carrasco, 2011; Corbetta & Shulman, 2002)
100 and previous studies indicate similarities between the neural circuits
101 implemented in orienting attention towards representations in perception and
102 memory (Awh & Jonides, 2001; Heuer & Schubö, 2016; Lepsien & Nobre, 2006;
103 Myers, Walther, Wallis, Stokes, & Nobre, 2015).

104 Altogether, it urges experimental designs with more sensitive task, as well
 105 as varied levels of task difficulty, to reveal the potential difference between the
 106 endogenous and exogenous attentional effects in the memory domain, in such
 107 a way we could unify the attentional effects in both perception and memory.
 108 Moreover, neuroimaging methods with high temporal resolutions are needed to
 109 catch the potential difference in neural processes between them. Therefore, we
 110 conducted 4 experiments, combined with physiological and neural
 111 measurements, including eye-tracker, electroencephalography (EEG),
 112 magnetoencephalography (MEG), and transcranial magnetic stimulation (TMS),
 113 to address the following two questions: 1) whether there is a behavioral or
 114 neural difference between the endogenous and exogenous attentional effects
 115 in the memory domain? 2) If there exists a difference, when and how the
 116 endogenous and exogenous attentional effects differentiate?



117
 118 **Figure 1 | Trial sequence.** a, Example of two types of valid retro-cue trials (endogenous
 119 and exogenous cue) and of a no cue trial. b, Configurations of the display that was set to
 120 load-2 in four experiments.

121

122 RESULTS

123 Experiment 1

124 In the first experiment, a total of 64 participants were requested to
 125 memorize two or four oriented Gabor patches (presented for 500 ms) that
 126 differed by at least 10 degrees. After a delay of 2 s, one of the Gabor patches
 127 was probed. The participants reproduced its orientation using a computer
 128 mouse as precisely as possible. In two thirds of all the experimental trials, a

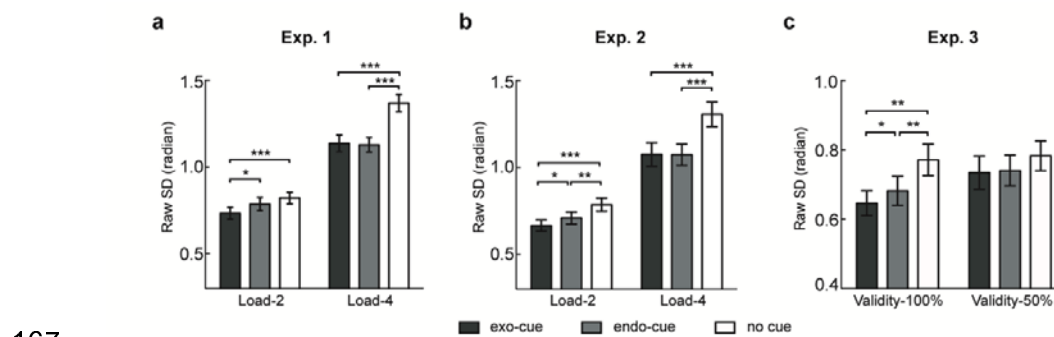
129 retro-cue was presented in the middle of the 2 s delay. Half of the retro-cue
130 trials used an endogenous cue (i.e. an arrow at the center of the screen) and
131 the other half used an exogenous cue (i.e. a dashed circle outside of one of the
132 Gabor patches). All cues were 100% valid, indicating the upcoming probed
133 location. All trials were randomly presented. To control eye movements and
134 blinks, gaze positions were on-line monitored during task performance (see
135 Methods for more details).

136 As the best fitting model of memory recall task were still in debate (Adam,
137 Vogel, & Awh, 2017; Ma, Husain, & Bays, 2014), we mainly focused on
138 reporting the results from the raw circular standard deviation of the recall error
139 (i.e. Raw SD). Six participants in experiment 1 were excluded from further
140 analysis due to either poor performance (lower end in the group 99%
141 confidence interval, the same below) in at least one condition (4 participants) or
142 data missing (2 participants). Raw SD from the rest of 58 subjects was shown in
143 Fig. 2a. Baseline-corrected pupillary diameters were calculated as an index to
144 reflect physiological states (e.g. larger pupil size is associated with a higher
145 level of arousal (Alnæs et al., 2014; Granholm & Steinhauer, 2004; Nassar et
146 al., 2012)) over time. The pupillary diameters are illustrated in Fig. 3a-b. We
147 performed two-way repeated measures ANOVAs with the within-subject factors
148 of LOAD (2 and 4) and ATTENTION (endogenous, exogenous, and no-cue) for
149 behavioral results. Post-hoc two-tailed t-tests were performed to examine the
150 effect of ATTENTION under each load, and multiple comparisons were
151 corrected using the false discovery rate (FDR) correction (the same for the
152 following experiments).

153 2-by-3 ANOVAs on the Raw SD revealed that there were significant main
154 effect of LOAD ($F(1, 57) = 369.47, p = 1.4e-26, \eta_p^2 = 0.87$), significant main
155 effect of ATTENTION ($F(2, 114) = 46.21, p = 2.0e-15, \eta_p^2 = 0.45$), and a
156 significant interaction between LOAD and ATTENTION ($F(2, 114) = 26.40, p$
157 $= 3.8e-10, \eta_p^2 = 0.32$).

158 Figure 2a displays the interaction. Under the load-4 condition, both
159 endogenous and exogenous retro-cues facilitated VWM performance
160 equivalently, compared to the no-cue condition (4_endo vs. 4_exo: $t(57) = -0.38,$
161 $p_{adj} = 0.71$, Cohen's $d = 0.05$; 4_endo vs. 4_neu: $t(57) = -8.20, p_{adj} = 4.8e-11,$

162 Cohen's $d=1.08$; 4_exo vs. 4_neu: $t(57)=-8.59$, $p_{adj}=2.2e-11$, Cohen's $d=1.13$.
 163 Importantly, under the load-2 condition, exogenous cues were more effective
 164 than endogenous cues (2_endo vs. 2_exo: $t(57)=2.64$, $p_{adj}=0.017$, Cohen's
 165 $d=0.35$; 2_endo vs. 2_neu: $t(57)=-1.70$, $p_{adj}=0.094$, Cohen's $d=0.22$; 2_exo vs.
 166 2_neu: $t(57)=-5.82$, $p_{adj}=8.5e-7$, Cohen's $d=0.76$).



167
 168 **Figure 2 | Behavioral results.** **a**, Behavioral results from experiment 1 ($n=58$). **b**,
 169 Behavioral results from experiment 2 ($n=19$). **c**, Behavioral results from experiment 3
 170 ($n=24$). * $p<0.05$, ** $p<0.01$, *** $p<0.001$, all p -values are FDR corrected.

171
 172 To focus on the attentional processes induced by the retro-cues, 2-by-2
 173 ANOVAs for the pupil diameters were performed at each time point. There
 174 existed significant main effects of LOAD starting from 284 ms after the sample
 175 stimuli until 376 ms after the cues ($ps<0.05$), significant main effects of
 176 ATTENTION (endogenous, exogenous) starting from 350 ms after the cues
 177 ($ps<0.05$), and significant interactions starting from 406 ms after the cues
 178 ($ps<0.05$). To further demonstrate this interaction, the sensory evoked activity
 179 could be subtracted between the two VWM loads within one attentional type.
 180 The difference between endogenous and exogenous attention was shown in
 181 Fig. 3b. Within the time window of post-cue 300-500ms, a larger pupillary
 182 dilation for exogenous attention was observed (endo (2-4) VS. exo (2-4):
 183 $t(57)=-2.29$, $p=0.026$, Cohen's $d=0.30$).

184 The results remained largely unchanged when we dropped out the trials
 185 without perfect central fixations (see Methods, supplementary Fig. S1a and Fig.
 186 S2a-b). These findings suggested that exogenous attention would induce a
 187 higher level of alertness which was usually associated with larger pupil dilation.

188

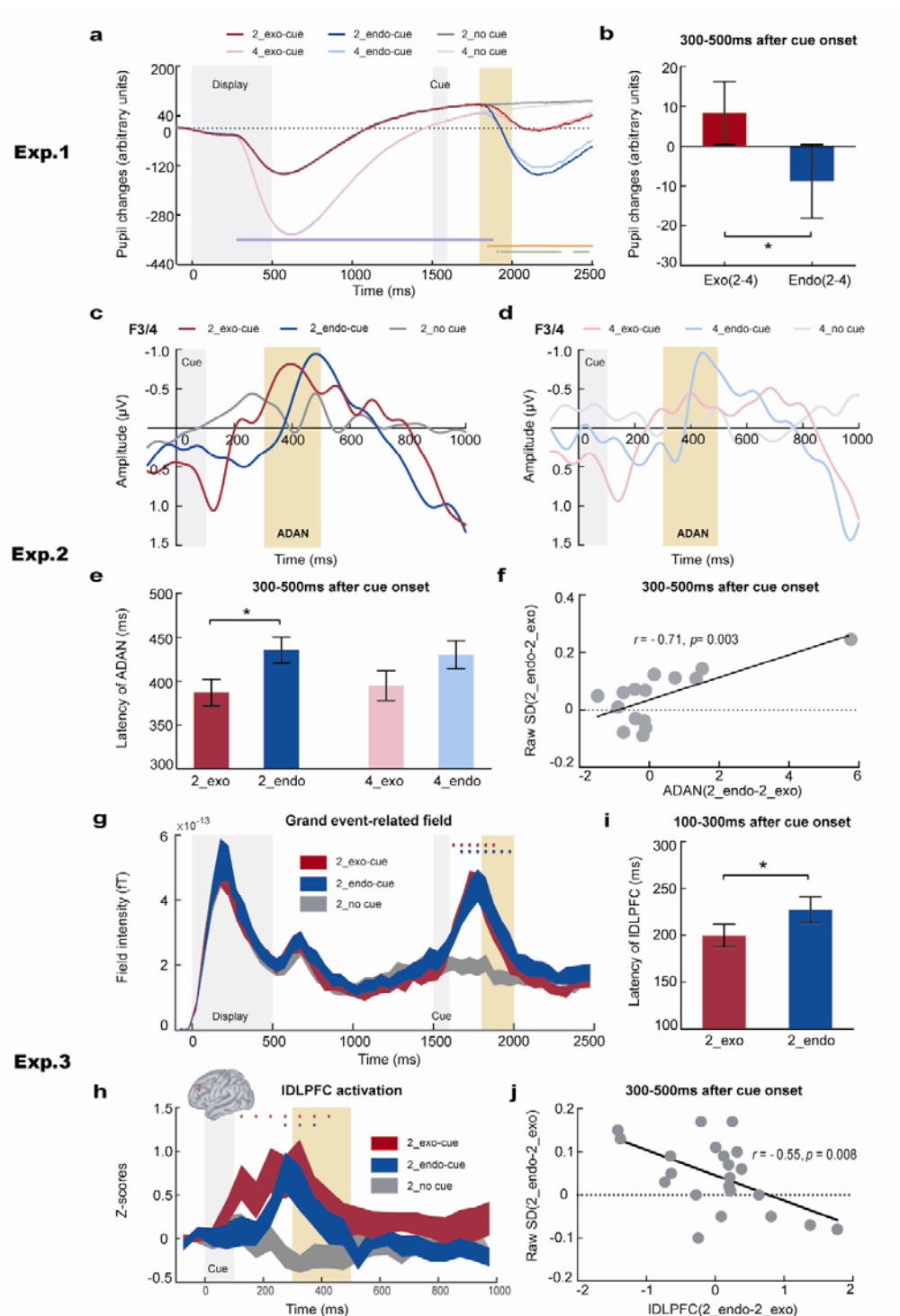


Figure 3 | Psychophysiological results. **a**, Pupil changes as a function of time referred to the memory display. Purple dots indicate a main effect of LOAD, orange dots indicate a main effect of ATTENTION, and green dots indicate a 2-by-2 interaction. **b**, Load-by-cue interactions within the time window of 300-500 ms after the cue onset. **c-d**, The contra-minus ipsilateral curves at frontal F3/4 electrodes time-locked to the retro-cue onset for the

load-2 (c) and load-4 condition (d). The yellow area indicates the time period of the ADAN component (300-500 ms after the cue onset). **e**, The latency of the ADAN component for retro-cue conditions. **f**, Correlations between ADAN and behavioral results under the load-2 condition (two-tailed *t*-tests, $p < 0.05$). **g-h**, Grand event-related field time-locked to memory display (g) and the left DLPFC activation time-locked to the cue onset (h) in blocks with 100% validity. Red (blue) dots indicate higher brain responses to exogenous (endogenous) retro-cues than to no cue condition (one-tailed *t*-tests, $p < 0.05$). Yellow areas indicate a time period of 300-500 ms after the cue onset. **i**, The latency of IDLPFC activation within the range of 100-300 ms after cue onset for retro-cue conditions. **j**, Correlations in blocks with 100% validity of activation differences between endogenous and exogenous cues at IDLPFC and their behavioral differences within a post-cue period of 300-500ms. * $p < 0.05$.

207

208

209 Experiment 2

210 To further identify the neural dynamics underlying the behavioral
211 distinction between effects of the two types of retro-cues, experiment 2 was
212 performed by another group of 20 participants with EEG. The design was
213 identical to that in experiment 1 with only one exception for the layout of Gabor
214 patches, where under the load-2 condition, the Gabor patches were bilaterally
215 presented on both hemispheres of the screen to balance the visual inputs. One
216 participant was excluded from behavioral analysis due to poor performance
217 (see the above criterion) and another 4 participants were excluded from EEG
218 analysis due to excessive artifacts in EEG signals. Again, behavioral results
219 demonstrated a significant main effect of LOAD ($F(1, 18) = 116.31$, $p = 2.8e-9$,
220 $\eta_p^2 = 0.87$), a significant main effect of ATTENTION ($F(2, 36) = 43.27$, p
221 $= 2.7e-10$, $\eta_p^2 = 0.71$), and a significant LOAD*ATTENTION interaction ($F(2, 36)$
222 $= 7.10$, $p = 0.003$, $\eta_p^2 = 0.28$). Post hoc *t* tests further revealed an advantage of
223 exogenous retro-cues only under the load-2 condition (Fig. 2b; 2_endo vs.
224 2_exo: $t(18) = 2.25$, $p_{adj} = 0.037$, Cohen's $d = 0.52$; 4_endo vs. 4_exo: $t(18) = 0.06$,
225 $p_{adj} = 0.952$, Cohen's $d = 0.01$), replicating the findings in experiment 1.
226 Furthermore, such benefit of the exogenous cue was not due to the
227 speed-accuracy trade-off effect, as no reaction time (RT) difference was

228 observed between the two types of retro-cues under any memory loads (see
229 Supplementary Fig. S1b).

230 One lateralized event-related potential (ERP) component, an anterior
231 directing attention negativity (ADAN, 300-500ms after cue onset) at frontal F3/4
232 electrodes, which was tightly related to attentional orientation (Eimer, Velzen, &
233 Driver, 2002; Göddertz, Klatt, Mertes, & Schneider, 2018; Myers et al., 2015),
234 showed a shorter latency for exogenous retro-cues than that for endogenous
235 retro-cues under the load-2 condition (Fig. 3c-e, see supplementary Fig. S2e-f
236 for un-subtracted bilateral waves; 2_endo vs. 2_exo: $t(14)=2.20$, $p=0.045$,
237 Cohen's $d=0.57$; 4_endo vs. 4_exo: $t(14)=1.50$, $p=0.156$, Cohen's $d=0.387$).
238 Interestingly, differences in ADAN amplitude were further positively correlated
239 with differences between the two types of cues under the load-2 condition in
240 Raw SD ($r=0.71$, $p=0.003$) (Fig. 3f). Although similar neural-behavioral
241 correlations were found for the posterior contralateral negativity (PCN,
242 supplementary Fig. S2), which is also associated with attentional orientation
243 (Eimer et al., 2002; Göddertz et al., 2018; Heuer & Schubö, 2016; Luck &
244 Hillyard, 1994; Myers et al., 2015), PCN failed to show the difference in latency
245 between the two types of retro-cues.

246 In addition, lateralized early ERPs, such as P1pc and P1ac (100-160ms
247 after the cue onset), were exclusively observed in exogenous retro-cue
248 conditions under both memory loads (supplementary Fig. S2k and S2o),
249 whereas the distractor positivity (Pd, another lateralized ERP, 580-680ms after
250 cue onset), which has been linked to the inhibition of attentional shift
251 (Schneider, Barth, Getzmann, & Wascher, 2017; Schneider, Mertes, &
252 Wascher, 2016), was solely identified in endogenous retro-cue conditions
253 (supplementary Fig. S2n).

254

255 Experiment 3

256 Although the EEG recording has the advantage of high temporal resolution,
257 it still lacks of spatial resolution, and is not well-performed in source analysis,
258 which could be remedied by MEG recording. Therefore, experiment 3 was
259 performed by another group of 27 participants in a magnetoencephalography
260 (MEG) scanner. As in the perceptual domain, when the cue validity drops to

261 50%, effects induced by endogenous cues diminish while effects induced by
262 exogenous cues remain (Briand, 1998; Kingstone, Smilek, Ristic, Friesen, &
263 Eastwood, 2003; Posner, 1978). To exclude cue validity as a potential factor
264 that could influence endogenous vs. exogenous cues differently, we added a
265 condition of cue validity with 50% in experiment 3. Meanwhile, in order to draw
266 similar number of trials as in experiments 1 and 2, we focused on the load-2
267 condition in which endogenous and exogenous cues induced behavioral
268 difference in the above experiments. The cue validity was block-wise designed
269 with 100% vs. 50%. The combination of retro-cue types (endogenous or
270 exogenous cue) and cue validities (100% or 50%) was presented at separate
271 blocks, randomly intermixed with no-cue trials (see Methods for details).

272 Two subjects were excluded due to poor performance. The behavioral
273 results revealed a significant main effect of ATTENTION when cues were 100%
274 valid ($F(2, 46) = 12.24, p = 5.5e-5, \eta_p^2 = 0.35$). Post hoc t tests further showed
275 that exogenous cues led to lower Raw SD compared to endogenous cues
276 (endo vs. exo: $t(23)=2.20, p_{adj}=0.038$, Cohen's $d=0.45$), consistent with results
277 in experiments 1 and 2. In contrast, non-significant effect of ATTENTION was
278 observed when cues were 50% valid (Fig. 2c; $F(2, 46) = 1.14, p = 0.329, \eta_p^2 =$
279 0.047), although valid cues resulted in better performance than invalid cues
280 (supplementary Fig.S1c; valid_endo vs. invalid_endo: $t(23)=-2.58, p=0.017$,
281 Cohen's $d=0.53$; valid_exo vs. invalid_exo: $t(23)=-3.46, p=0.002$, Cohen's
282 $d=0.71$). These results were different from the traditional findings in the
283 perceptual domain, in which exogenous cues always attracted attention even
284 when they were uninformative (Briand, 1998; Carrasco, 2011; Emmanuel
285 Guzman, Marcia Grabowecky, German Palafox, 2012; Kingstone et al., 2003;
286 Posner, 1978), suggesting that the differences between the two types of
287 retro-cues in our study was not due to automatic attentional processing. Again,
288 the control analysis on RTs in experiment 3 excluded the possibility of
289 speed-accuracy trade-off (supplementary Fig. S1d). Taken together, there was
290 a consistent behavioral difference between endogenous and exogenous
291 retro-cues with 100% validity under the load-2 condition.

292 Another two subjects were excluded for MEG analysis due to large
293 artifacts in MEG signals (see Methods for details). To assess brain areas that

might lead to behavioral changes, we estimated neural responses in the source space for each condition. Two clusters including the right intraparietal sulcus (IPS) and left DLPFC displayed significantly larger neural activation in both cued conditions vs. no-cue condition in the post-cue period of 0-500 ms (permutation test $\alpha < 0.05$). During the post-cue period of 100-300ms, activations at the left DLPFC (lDLPFC) showed a shorter latency for the exogenous cue than that for the endogenous cue (Fig. 3h-i; endo vs. exo: $t(21)=2.20$, $p=0.040$, Cohen's $d=0.47$). Furthermore, during the post-cue period of 300-500ms which overlapping the time window of ADAN, differences in activation between endogenous and exogenous cues at the lDLPFC were negatively correlated with their behavioral differences in Raw SD ($r=-0.55$, $p=0.008$) (Fig. 3j). These findings, together with the ADAN in experiment 2 (Fig. 3g), further suggested that although in the perception domain, exogenous cues were regarded as merely bottom-up manner (Buschman & Miller, 2007; Hahn et al., 2006; Theeuwes, 2010), they induced a faster and stronger lDLPFC activation, which accounted for better performance than that induced by endogenous retro-cues. In addition, we found that activations at right IPS (rIPS) during the post-cue 100-300ms was marginally faster for exogenous than endogenous cues (supplementary Fig. S3a-b; endo vs. exo: $t(21)=2.06$, $p=0.052$, Cohen's $d=0.44$), and could predict both retro-cue benefits in Raw SD (supplementary Fig. S3c-d; endo-neu: $r=-0.55$, $p=0.008$; exo-neu: $r=-0.60$, $p=0.003$), suggesting shared neural mechanisms by the two types of retro-cues in the parietal cortex.

Indeed, the superior parietal lobe (SPL) vs. the inferior parietal lobe (IPL) have been regarded as critical regions in top-down vs. bottom-up attention-related processes (Ciaramelli, Grady, & Moscovitch, 2008; Corbetta, Kincade, & Shulman, 2002; Corbetta & Shulman, 2002). To understand how visual information was transferred across parietal and frontal regions and how such process was modulated by frontal activities relating to behavior, Granger causality analysis was performed to the source activity of SPL, IPL, DLPFC, and the lateral occipital cortex (LOC) (see Methods for details, supplementary Fig. S3e). Results revealed that both endogenous and exogenous cues induced a top-down control from DLPFC to LOC during 300-500ms after the

cue onset (Fig. 4a-b). The time window coincided with that of the frontal-behavioral correlation in experiment 3 (Fig. 3j), the pupil size difference in experiment 1 (Fig. 3b) and ADAN component in experiment 2 (Fig. 3f). Interestingly, such process exclusively took place as early as 200ms after the onset of an exogenous cue, consistent with the time window of IDLPFC activations in experiment 3 (Fig. 3h-i) and P1ac in experiment 2 (supplementary Fig. S2o). These findings further suggested that exogenous retro-cues induced early activation in the frontal-posterior network, which was quite different from the merely bottom-up processes induced by exogenous cues in perception (Buschman & Miller, 2007; Chica et al., 2013; Connor, Egeth, & Yantis, 2004; Dugué, Merriam, Heeger, & Carrasco, 2018; Eimer et al., 2002; Hickey, Van Zoest, & Theeuwes, 2010; Hopfinger & West, 2006; Theeuwes, 2010). For the no-cue condition, there was no information transfer between the frontal and occipital regions (Fig. 4c).

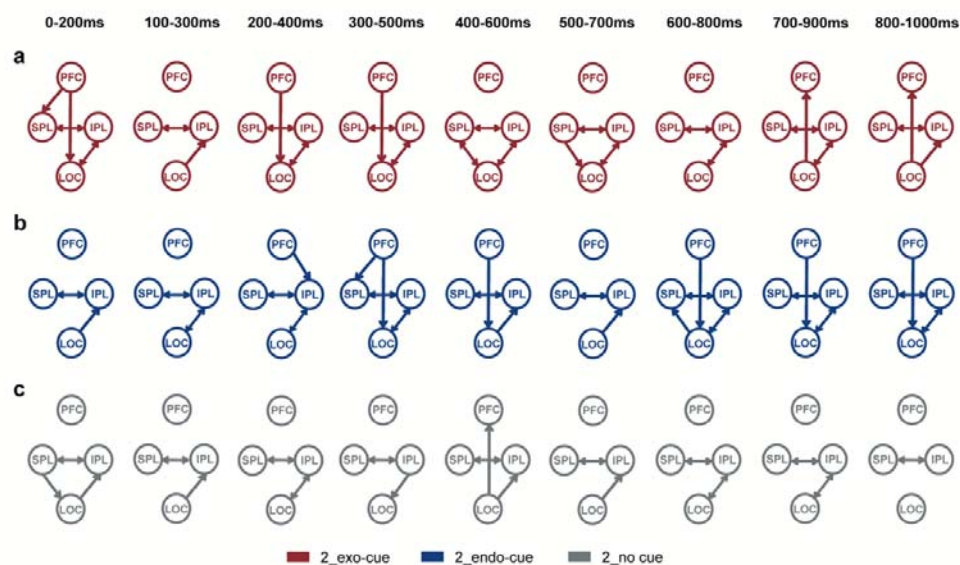


Figure 4 | Granger causality analysis from experiment 3. a-b, Neural dynamics for exogenous (a) or endogenous (b) retro-cue trials with 100% validity, time-locked to the cue onset. **c,** Neural dynamics for no-cue trials in blocks with 100% validity, time-locked to the middle of delay period. Each arrow indicates the direction of information flow after Bonferroni correction ($p < 0.05$).

Experiment 4

349 Finally, we continued to identify the causal role of DLPFC and IPS in the
350 retro-cue effects in experiment 4 with transcranial magnetic stimulation (TMS).
351 A new group of 22 subjects were recruited. In each block, three types of cues
352 were randomly mixed with equal number of trials. Stimulating sites were
353 targeted at IDLPFC or rIPS, whose MNI coordinates were acquired as the
354 maximal activation observed in experiment 3. Vertex served as a control
355 stimulating site. To explore the time course of these brain regions in cognitive
356 processes, a single pulse TMS was given at 100, 400, or 700ms after the cue
357 onset (when no cue presented, we match the time points with the cued
358 condition) in each trial. After considering the length of the experiment and
359 maximum number of stimulating pulses in one session, we focused on the
360 load-2 condition with 100% validity, in which all three experiments showed
361 behavioral difference between endogenous and exogenous cues (see Methods
362 for details).

363 Two subjects were excluded due to poor performance. When TMS was
364 applied at vertex (Fig. 5a), there was a main effect of ATTENTION ($F(2, 38) =$
365 $18.11, p = 3.0e-6, \eta_p^2 = 0.49$). Post hoc *t*-tests replicated the findings in the prior
366 three experiments, indicating a superiority effect of exogenous retro-cues
367 (endo vs. exo: $t(19) = 2.33; p_{adj} = 0.031$; Cohen's $d = 0.52$). Interestingly, TMS
368 over IDLPFC or rIPS both abolished the differences in Raw SD between the two
369 types of retro-cues (IDLPFC: $t(19) = -1.62, p_{adj} = 0.121$, Cohen's $d = 0.36$; rIPS:
370 $t(19) = 0.963, p_{adj} = 0.348$, Cohen's $d = 0.22$).

371 However, further looking at TMS results at different time points, they
372 further suggested the timing of IDLPFC/rIPS function. When TMS was
373 delivered at 100ms after the exogenous cue onset, the retro-cue benefit in Raw
374 SD diminished (Fig. 5b; IDLPFC_100ms vs. vertex: $t(19) = 2.80, p = 0.012$,
375 Cohen's $d = 0.62$). The control analysis on endogenous retro-cue effects at any
376 stimulating time points did not replicate such results, neither did rIPS-targeted
377 TMS at both types of retro-cues (Fig. 5b-c). These findings strongly proved that
378 IDLPFC (but not rIPS) causally affected exogenous (but not endogenous)
379 attentional processing within a time window of 100ms.

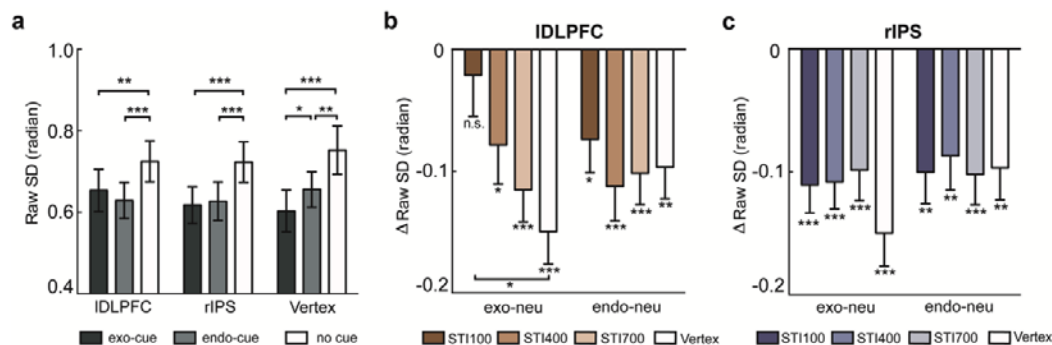


Figure 5 | TMS results from experiment 4 (n=20). **a**, Merged data when TMS was set to LPFC, RIPS and vertex, respectively. **b-c**, Retro-cue effects when IDLPFC-targeted (**b**) or rIPS-targeted (**c**) TMS was delivered at 100ms, 400ms, and 700ms after the cue onset. 'exo' stands for exogenous cue, 'endo' stands for endogenous cue, and 'neu' stands for no cue condition. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, all p -values are FDR corrected.

DISCUSSION

Consistent with previous findings using retro-cues under the VWM load-4 condition (Kuo, Stokes, & Nobre, 2012; Murray et al., 2013; Pertzov et al., 2013; Pertzov, Manohar, & Husain, 2017; Shimi et al., 2014; Astle, Summerfield, Griffin, & Nobre, 2012; Gözenman, Tanoue, Metoyer, & Berryhill, 2014; Gunseli, van Moorselaar, Meeter, & Olivers, 2015; Lepsien & Nobre, 2006; Griffin & Nobre, 2003; Tanoue & Berryhill, 2012), we observed equivalent behavioral benefit for both exogenous and endogenous cues, confirming the important roles of spatial attention in modulating working memory representations. More importantly, we discovered a reliable difference in behavior between the two attentional effects under the low (but not high) VWM load through four experiments with 133 participants, which was not investigated in the previous studies (Kuo et al., 2012; Murray et al., 2013; Pertzov et al., 2013, 2017; Shimi et al., 2014; Astle, Summerfield, Griffin, & Nobre, 2012; Brady & Hampton, 2018; Gözenman, Tanoue, Metoyer, & Berryhill, 2014; Gunseli, van Moorselaar, Meeter, & Olivers, 2015; Lepsien & Nobre, 2006; Williams, Hong, Kang, Carlisle, & Woodman, 2013; Griffin & Nobre, 2003; Matsukura, Cosman, Roper, Vatterott, & Vecera, 2014; Tanoue & Berryhill, 2012). Combining multimodal neurophysiological recordings with eye-tracking, EEG, MEG and TMS, we

406 further revealed convergent evidence for physiological and neural differences
407 between the two attentional processes under the low VWM load.

408 Specifically, compared to endogenous cues, exogenous cues led to
409 shorter latency of frontal ERP component (i.e. ADAN) indicating faster
410 attentional deployment (experiment 2), and earlier DLPFC activity that were
411 normally regarded as high-level processes of cognitive functions (experiment
412 3). Importantly, both amplitude of ADAN (experiment 2) and activity in IDLPFC
413 (experiment 3) from 300ms to 500ms after the retro-cue onset explained the
414 behavioral difference between the two types of cues.

415 How does such faster frontal activity in exogenous attention than
416 endogenous attention come out and lead to better performance? Our data
417 demonstrate one possibility. In experiment 2, there existed early contralateral
418 ERP component (i.e. P1ac and P1pc) between 100-200ms after the exogenous
419 cue, indicating fast automatic attentional processes, which however, did not
420 exist in the endogenous condition. Such time window coincided with the time
421 window in experiment 3 when the exogenous cues induced earlier DLPFC and
422 IPS activity. Interestingly, the pupil dilation, which was tightly connected to the
423 arousal systems (Alnæs et al., 2014; Granholm & Steinhauer, 2004; Nassar et
424 al., 2012), appeared stronger for the exogenous cues than the endogenous
425 cues at a similar latency of around 400ms (experiment 1). The overlapped time
426 window of 300-500ms might indicate interactions between the alerting and
427 orienting attention networks (Fan et al., 2005; Gazzaley & Nobre, 2012). To
428 sum up, through the fast-automatic bottom-up processing, exogenous cues
429 recruit early top-down control from DLPFC, and induce earlier frontal activities,
430 which further interact with the arousal system that may amplify the processes.
431 Such frontal activity induced by the exogenous cue is the real reason for the
432 larger behavioral benefit than the endogenous cue.

433 However, it should be noted that both DLPFC and IPS demonstrate earlier
434 activity in exogenous condition (experiment 3), and both the frontal ADAN and
435 the parietal PCN could explain the behavioral difference between the two
436 attentional types (experiment 2). Meanwhile, there are lots of evidence that
437 prefrontal cortical areas receive bottom-up projection from parietal and occipital
438 areas (Bressler, Tang, Sylvester, Shulman, & Corbetta, 2008; Connor et al.,

439 2004; Theeuwes, 2010). Then which activity is more critical, the frontal or the
440 parietal? In the current task-set, we suggested the more important role in the
441 frontal top-down modulation, as TMS over DLPFC abolished the exogenous
442 superiority but TMS over IPS did not (experiment 4). Nevertheless, it should be
443 acknowledged that there might be other parietal areas that contribute to these
444 processes as we only tested IPS in the present study.

445 Alternative explanations may exist for such top-down signal from DLPFC.
446 For example, DLPFC has been suggested to maintain VWM representations in
447 a plethora of studies (Barbey, Koenigs, & Grafman, 2013; Curtis & D'Esposito,
448 2003; S. Funahashi, Bruce, & Goldman-Rakic, 1989; Shintaro Funahashi,
449 Chafee, & Goldman-Rakic, 1993; Shintaro Funahashi, Takeda, & Watanabe,
450 2004; Fuster, 1971). Recent studies reveal that the representations maintained
451 in DLPFC neurons could be flexibly switched during the delay (Spaak,
452 Watanabe, Funahashi, & Stokes, 2017; Stokes et al., 2013). Therefore, the
453 top-down signal observed from the present study could also be representations
454 stored in DLPFC and exogenous cues lead to faster shift of tuning of the
455 representations (Spaak et al., 2017; Stokes et al., 2013). Indeed, dynamic
456 coding in the prefrontal cortex has recently been regarded as the format of
457 VWM storage (Meyers, 2018; Stokes, 2015), in contrast to the traditional view
458 of static sustained activity (Goldman-Rakic, 1995). Our data put another
459 possibility that attentional processes might also play a role in such dynamic
460 shifting of representations.

461 Moreover, our data could unify attentional processes in the perceptual
462 domain and those in the memory domain on the one side. In the current study,
463 inspired by findings from the perceptual domain that the exogenous and
464 endogenous attention effects tend to differentiate when sensory signal-to-noise
465 ratio is high (Lu & Doshier, 1998, 2000; Lu, Lesmes, & Doshier, 2002), we
466 employ a low load condition in which memory signal-to-noise ratio is high, as
467 well as more sensitive measures with the memory recall task compared to the
468 prior change detection task. Interestingly, we consistently find behavioral
469 difference between exogenous and endogenous cues in four experiments. In
470 such a way, the attentional effects in the perception domain and the memory
471 domain can be incorporated by the same criteria of signal-to-noise ratio:

472 endogenous and exogenous cues are equivalent when the signal-to-noise ratio
473 is low; they tend to differentiate when the signal-to-noise ratio is high. However,
474 future studies are needed to quantify the boundary of signal-to-noise ratio
475 leading to the difference.

476 Meanwhile, at the neural level, contralateral ERP components indicating
477 the attentional shift, such as ADAN and PCN, were identified for both endo- and
478 exogenous retro-cues in the current study (experiment 2). Similar results were
479 found in those EEG studies using spatial pre-cues (Eimer et al., 2002; Heuer &
480 Schubö, 2016; Hickey, Lollo, & McDonald, 2009; Luck & Hillyard, 1994). Both
481 endogenous and exogenous cues activated a large network involving parietal,
482 frontal and visual cortices (experiment 3), overlapping the areas in fMRI studies
483 using the pre-cues (Corbetta, Kincade, Ollinger, Mcavoy, & Gordon, 2000;
484 Lepsien & Nobre, 2006; Mao, Zhou, Zhou, & Han, 2006; Rosen et al., 1992;
485 Vandenberghe, Gitelman, Parrish, & Mesulam, 2001; Yantis et al., 2002;
486 Dugué et al., 2018). These findings altogether indicated that spatial orienting in
487 both the perceptual and VWM domain shared some common neural
488 substrates.

489 On the other side, however, the attentional processes in perception or
490 memory can be fairly different. Traditionally, exogenous attentional processes
491 in perception were regarded as a merely bottom-up manner (Buschman &
492 Miller, 2007; Chica et al., 2013; Connor et al., 2004; Dugué et al., 2018; Eimer
493 et al., 2002; Hickey et al., 2010; Hopfinger & West, 2006; Theeuwes, 2010).
494 However, in the present study, Granger causality analysis with source activity
495 in MEG suggested a top-down signal from the prefrontal cortex as early as
496 100ms after the exogenous cue (experiment 3), and single pulse TMS over
497 IDLPFC at the same 100ms post-cue time indeed abolished the exogenous
498 retro-cue effects (experiment 4). These results indicated that 1) exogenous
499 attention processes in memory were not only bottom-up; 2) the early top-down
500 influence from the prefrontal cortex was critical for the superiority of exogenous
501 attentional effects.

502 Besides the early top-down signal distinguishing exogenous attention
503 processes in the memory domain from those in the perceptual domain, there
504 still existed other differences between the attentional processes in the two

domains, based on our results. For example, exogenous cues behaved differently in the present study from that in the perceptual domain when the cue validity was 50%. Previous studies suggested that in the perceptual domain exogenous cues always attracted attention in the condition with 50% validity while endogenous cues did not (Briand, 1998; Kingstone et al., 2003; Posner, 1978). However, in the current results, both of their retro-cue benefits had gone when the cues were 50% valid, although differences between valid and invalid trials still existed. Altogether, at least we could say the attentional processes in memory vs. perception are similar but not the same, and the endogenous vs. exogenous attention processes in memory are different from the traditional top-down vs. bottom-up dissection in perception.

As predicted from the perceptual domain, the presence of superiority of exogenous cues towards VWM representations may be generalized to long term memory (LTM). It is of great interest to further demonstrate the two different types of attention effects on LTM at different signal-to-noise ratio of memory trace. Such findings may also facilitate the educational implications that when the task is easier and memory signal is strong, exogenous cues might be more efficient than endogenous cues.

Although we combined different research modalities to form convergent evidence to strengthen, and further to explain the superiority of exogenous attention on VWM when the memory load was low, there still remained questions to be resolved. For example, after the exogenous cue presents, how is the cue processed in sensory cortices and then conveyed to the prefrontal cortex to trigger the top-down control signal? Or it may be conveyed directly from subcortical areas such as the superior colliculus or pulvinar, given such a fast time scale. To addressing these questions may require invasive neurophysiological studies, combining with computational models. We look forward to future studies using these tools to fully understand the interaction between attention and VWM, both lies in the center of human cognition.

REFERENCES

536 Adam, K. C. S., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working
537 memory. *Cognitive Psychology*, 97, 79–97.
538 <https://doi.org/doi.org/10.1016/j.cogpsych.2017.07.001>

539 Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B.
540 (2014). Pupil size signals mental effort deployed during multiple object tracking and
541 predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of*
542 *Vision*, 14(4), 1–20. <https://doi.org/doi:10.1167/14.4.1>

543 Astle, D. E., Summerfield, J., Griffin, I., & Nobre, A. C. (2012). Orienting attention to locations in
544 mental representations. *Attention, Perception, & Psychophysics*, 74(1), 146–162.
545 <https://doi.org/doi.org/10.3758/s13414-011-0218-3>

546 Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working
547 memory. *Trends in Cognitive Sciences*, 5(3), 119–126.
548 [https://doi.org/10.1016/S1364-6613\(00\)01593-X](https://doi.org/10.1016/S1364-6613(00)01593-X)

549 Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

550 Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to
551 human working memory. *Cortex*, 49(5), 1195–1205.
552 <https://doi.org/10.1016/j.cortex.2012.05.022>

553 Barnett, L., & Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: a new
554 approach to Granger-causal inference. *Journal of Neuroscience Methods*, 223(15), 50–68.
555 <https://doi.org/doi.org/10.1016/j.jneumeth.2013.10.018>

556 Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical*
557 *Software*, 31(10), 34019. <https://doi.org/10.18637/jss.v031.i10>

558 Brady, R. J., & Hampton, R. R. (2018). Post-encoding control of working memory enhances
559 processing of relevant information in rhesus monkeys (*Macaca mulatta*). *Cognition*, 175,
560 26–35. <https://doi.org/10.1016/j.cognition.2018.02.012>

561 Bressler, S. L., Tang, W., Sylvester, C. M., Shulman, G. L., & Corbetta, M. (2008). Top-Down
562 Control of Human Visual Cortex by Frontal and Parietal Cortex in Anticipatory Visual
563 Spatial Attention. *Journal of Neuroscience*, 28(40), 10056–10061.
564 <https://doi.org/10.1523/JNEUROSCI.1776-08.2008>

565 Briand, K. A. (1998). Feature Integration and Spatial Attention: More Evidence of a Dissociation
566 between Endogenous and Exogenous Orienting. *Journal of Experimental Psychology*
567 *Human Perception & Performance*, 24(4), 1243–1256.
568 <https://doi.org/10.1037/0096-1523.24.4.1243>

569 Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the
570 prefrontal and posterior parietal cortices. *Science*, 315(5820), 1860–1864.
571 <https://doi.org/10.1126/science.1138071>

572 Busse, L., Katzner, S., & Treue, S. (2008). Temporal dynamics of neuronal modulation during
573 exogenous and endogenous shifts of visual attention in macaque area MT. *Proceedings*
574 *of the National Academy of Sciences*, 105(42), 16380–16385.
575 <https://doi.org/10.1073/pnas.0707369105>

576 Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Research*, 51(13), 1484–1525.
577 <https://doi.org/10.1016/j.visres.2011.04.012>

578 Chica, A. B., Bartolomeo, P., & Lupiáñez, J. (2013). Two cognitive and neural systems for
579 endogenous and exogenous spatial attention. *Behavioural Brain Research*, 237(1),
580 107–123. <https://doi.org/10.1016/j.bbr.2012.09.027>

581 Ciamarelli, E., Grady, C. L., & Moscovitch, M. (2008). Top-down and bottom-up attention to
582 memory: a hypothesis (AtoM) on the role of the posterior parietal cortex in memory
583 retrieval. *Neuropsychologia*, 46(7), 1828–1851.
584 <https://doi.org/10.1016/j.neuropsychologia.2008.03.022>

585 Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down.
586 *Current Biology*, 14(19), 850–852. <https://doi.org/10.1016/j.cub.2004.09.041>

587 Corbetta, M., Kincade, J. M., Ollinger, J. M., Mcavoy, M. P., & Gordon, L. (2000). Voluntary
588 orienting is dissociated from target detection in human posterior parietal cortex. *Nature*
589 *Neuroscience*, 3(6), 292–297. <https://doi.org/10.1038/73009>

590 Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and
591 their relationships to spatial working memory. *Journal of Cognitive Neuroscience*, 14(3),
592 508–523. <https://doi.org/10.1162/089892902317362029>

593 Corbetta, M., & Shulman, G. L. (2002). Control of Goal-Directed and Stimulus-Driven Attention
594 in the Brain. *Nature Reviews Neuroscience*, 3(3), 201–215.
595 <https://doi.org/10.1038/nrn755>

596 Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental
597 storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114.
598 <https://doi.org/10.1017/s0140525x01003922>

599 Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic*
600 *Bulletin and Review*, 24(4), 1158–1170. <https://doi.org/10.3758/s13423-016-1191-6>

601 Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working
602 memory. *Trends in Cognitive Sciences*, 7(9), 415–423.
603 [https://doi.org/10.1016/S1364-6613\(03\)00197-9](https://doi.org/10.1016/S1364-6613(03)00197-9)

604 Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial
605 EEG dynamics including independent component analysis. *Journal of Neuroscience*
606 *Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>

607 Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., ... Hyman, B.
608 T. (2006). An automated labeling system for subdividing the human cerebral cortex on
609 MRI scans into gyral based regions of interest. *NeuroImage*, 31(3), 968–980.
610 <https://doi.org/10.1016/j.neuroimage.2006.01.021>

611 Dugué, L., Merriam, E. P., Heeger, D. J., & Carrasco, M. (2018). Endogenous and exogenous
612 attention distinctly modulate fMRI activity in visual cortex. *BioRxiv*, 1–28.
613 <https://doi.org/10.1101/414508>

614 Eimer, M., Velzen, J. van, & Driver, J. (2002). Cross-Modal Interactions between Audition,
615 Touch, and Vision in Endogenous Spatial Attention: ERP Evidence on Preparatory States
616 and Sensory Modulations. *Journal of Cognitive Neuroscience*, 14(2), 254–271.
617 <https://doi.org/10.1162/089892902317236885>

618 Emmanuel Guzman, Marcia Grabowecky, German Palafox, S. S. (2011). A unique role of
619 endogenous visual-spatial attention in rapid processing of multiple targets. *Journal of*
620 *Experimental Psychology: Human Perception and Performance*, 37(4), 1065–1073.
621 <https://doi.org/10.1037/a0023514>

622 Engle, R., Kane, M., & W Tuholski, S. (1999). Individual differences in working memory
623 capacity and what they tell us about controlled attention, general fluid intelligence, and
624 functions of the prefrontal cortex. In *Models of Working Memory: Mechanisms of Active*
625 *Maintenance and Executive Control* (pp. 102–134). Cambridge: Cambridge University
626 Press. <https://doi.org/10.1017/CBO9781139174909.007>

627 Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation
628 of attentional networks. *NeuroImage*, 26(2), 471–479.
629 <https://doi.org/10.1016/j.neuroimage.2005.02.004>

630 Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic Coding of Visual Space
631 in the Monkey ' s Dorsolateral Prefrontal Cortex. *Journal of Neurophysiology*, 61(2),
632 331–349. <https://doi.org/10.1152/jn.1989.61.2.331>

633 Funahashi, Shintaro, Chafee, M. V, & Goldman-Rakic, P. S. (1993). Prefrontal neuronal activity
634 in rhesus monkeys performing a delayed anti-saccade task. *Nature*, 365(6448), 753–756.
635 <https://doi.org/10.1038/365753a0>

636 Funahashi, Shintaro, Takeda, K., & Watanabe, Y. (2004). Neural mechanisms of spatial
637 working memory □: Contributions of the dorsolateral prefrontal cortex and the thalamic
638 mediodorsal nucleus. *Cognitive, Affective, & Behavioral Neuroscience*, 4(4), 409–420.
639 <https://doi.org/10.3758/CABN.4.4.409>

640 Fuster, J. M. (1971). Neuron Activity Related to Short-Term Memory. *Science*, 173(3997),
641 652–654. <https://doi.org/10.1126/science.173.3997.652>

642 Gazzaley, A., & Nobre, A. C. (2012). The Attention System of the Human Brain: 20 Years After.
643 *Annual Review of Neuroscience*, 21(35), 73–89.
644 <https://doi.org/10.1146/annurev-neuro-062111-150525>.The

645 Göddertz, A., Klatt, L.-I., Mertes, C., & Schneider, D. (2018). Retroactive attentional shifts
646 predict performance in a working memory task: Evidence by lateralized EEG patterns.
647 *Frontiers in Human Neuroscience*, 12, 428. <https://doi.org/10.3389/fnhum.2018.00428>

648 Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, 14(3), 477–485.
649 [https://doi.org/https://doi.org/10.1016/0896-6273\(95\)90304-6](https://doi.org/https://doi.org/10.1016/0896-6273(95)90304-6)

650 Gözenman, F., Tanoue, R. T., Metoyer, T., & Berryhill, M. E. (2014). Invalid retro-cues can
651 eliminate the retro-cue benefit: Evidence for a hybridized account. *Journal of*
652 *Experimental Psychology □: Human Perception and Performance*, 40(5), 1748–1754.
653 <https://doi.org/10.1037/a0037474>

654 Granholm, E., & Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional
655 processes. *International Journal of Psychophysiology*, 52(1), 1–6.
656 <https://doi.org/10.1016/j.ijpsycho.2003.12.001>

657 Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations.
658 *Journal of Cognitive Neuroscience*, 15(8), 1176–1194.
659 <https://doi.org/10.1162/089892903322598139>

660 Griffin, I. C., & Nobre, A. C. (2012). Orienting attention to locations in internal representations.
661 *Journal of Cognitive Neuroscience*, 15(8), 1176–1194.
662 <https://doi.org/10.1162/089892903322598139>

663 Gunseli, E., van Moorselaar, D., Meeter, M., & Olivers, C. N. L. (2015). The reliability of
664 retro-cues determines the fate of noncued visual working memory representations.

665 *Psychonomic Bulletin & Review*, 22(5), 1334–1341.
666 <https://doi.org/10.3758/s13423-014-0796-x>

667 Hahn, B., Ross, T. J., & Stein, E. A. (2006). Neuroanatomical dissociation between bottom–up
668 and top–down processes of visuospatial selective attention. *NeuroImage*, 32(2), 842–853.
669 <https://doi.org/10.1016/j.neuroimage.2006.04.177>

670 Heuer, A., & Schubö, A. (2016). The focus of attention in visual working memory: Protection of
671 focused representations and its individual variation. *PLoS ONE*, 11(4), e0154228.
672 <https://doi.org/10.1371/journal.pone.0154228>

673 Hickey, C., Lollo, V. Di, & McDonald, J. J. (2009). Electrophysiological Indices of Target and
674 Distractor Processing in Visual Search. *Journal of Cognitive Neuroscience*, 21(4),
675 760–775. <https://doi.org/10.1162/jocn.2009.21039>

676 Hickey, C., Van Zoest, W., & Theeuwes, J. (2010). The time course of exogenous and
677 endogenous control of covert attention. *Experimental Brain Research*, 201(4), 789–796.
678 <https://doi.org/10.1007/s00221-009-2094-9>

679 Hopfinger, J. B., & West, V. M. (2006). Interactions between endogenous and exogenous
680 attention on cortical visual processing. *NeuroImage*, 31(2), 774–789.
681 <https://doi.org/10.1016/j.neuroimage.2005.12.049>

682 James, W., Burkhardt, F., Bowers, F., & Skrupskelis, I. K. (1890). *The principles of psychology*
683 (Vol. 1). Macmillan London.

684 Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, 37,
685 43–75. <https://doi.org/10.1146/annurev.ps.37.020186.000355>

686 Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention,
687 Researchers! It Is Time to Take a Look at the Real World. *Current Directions in*
688 *Psychological Science*, 12(5), 176–180. <https://doi.org/10.1111/1467-8721.01255>

689 Kuo, B. C., Stokes, M. G., & Nobre, A. C. (2012). Attention modulates maintenance of
690 representations in visual short-term memory. *Journal of Cognitive Neuroscience*, 24(1),
691 51–60. https://doi.org/10.1162/jocn_a_00087

692 Lepsien, J., & Nobre, A. C. (2006). Cognitive control of attention in the human brain: Insights
693 from orienting attention to mental representations. *Brain Research*, 1105(1), 20–31.
694 <https://doi.org/10.1016/j.brainres.2006.03.033>

695 Liu, T., Stevens, S. T., & Carrasco, M. (2007). Comparing the time course and efficacy of
696 spatial and feature-based attention. *Vision Research*, 47(1), 108–113.
697 <https://doi.org/10.1016/j.visres.2006.09.017>

698 Lu, Z.-L., & Doshier, B. A. (1998). External noise distinguishes mechanisms of attention. *Vision*
699 *Research*, 38(9), 1183–1198. <https://doi.org/10.1016/B978-012375731-9/50078-1>

700 Lu, Z.-L., & Doshier, B. A. (2000). Spatial attention: Different mechanisms for central and
701 peripheral temporal precues? *Journal of Experimental Psychology: Human Perception*
702 *and Performance*, 26(5), 1534–1548. <https://doi.org/10.1037/0096-1523.26.5.1534>

703 Lu, Z.-L., Lesmes, L. A., & Doshier, B. A. (2002). Spatial attention excludes external noise at the
704 target location. *Journal of Vision*, 2, 312–323. <https://doi.org/10.1167/2.4.4>

705 Luck, S. J., & Hillyard, S. (1994). Electrophysiological correlates of feature analysis during
706 visual search. *Psychophysiology*, 31, 291–308.
707 <https://doi.org/10.1111/j.1469-8986.1994.tb02218.x>

708 Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature*
709 *Neuroscience*, 17(3), 347–356. <https://doi.org/10.1038/nn.3655>

710 Mao, L., Zhou, B., Zhou, W., & Han, S. (2006). Neural correlates of covert orienting of visual
711 spatial attention along vertical and horizontal dimensions. *Brain Research*, 1136,
712 142–153. <https://doi.org/10.1016/j.brainres.2006.12.031>

713 Matsukura, M., Cosman, J. D., Roper, Z. J., Vatterott, D. B., & Vecera, S. P. (2014).
714 Location-specific effects of attention during visual short-term memory maintenance.
715 *Journal of Experimental Psychology: Human Perception and Performance*, 40(3),
716 1103–1116. <https://doi.org/10.1037/a0035685>

717 Meyers, X. E. M. (2018). Working Memory □: Neural Mechanisms Dynamic population coding
718 and its relationship to working memory. *Journal of Neurophysiology*, 120, 2260–2268.
719 <https://doi.org/10.1152/jn.00225.2018>

720 Miller, G. A. (1956). The Magical Number Seven , Plus or Minus Two Some Limits on Our
721 Capacity for Processing Information. *Psychological Review*, 63(2), 81–97.
722 <https://doi.org/10.1037/h0043158>

723 Murray, A. M., Nobre, A. C., Clark, I. A., Cravo, A. M., & Stokes, M. G. (2013). Attention
724 Restores Discrete Items to Visual Short-Term Memory. *Psychological Science*, 24(4),
725 550–556. <https://doi.org/10.1177/0956797612457782>

726 Myers, N. E., Walther, L., Wallis, G., Stokes, M. G., & Nobre, A. C. (2015). Temporal dynamics
727 of attention during encoding versus maintenance of working memory: complementary
728 views from event-related potentials and alpha-band oscillations. *Journal of Cognitive*
729 *Neuroscience*, 27(3), 492–508. https://doi.org/10.1162/jocn_a_00727

730 Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012).
731 Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature*
732 *Neuroscience*, 15(7), 1040–1046. <https://doi.org/10.1038/nn.3130>
733 Pertzov, Y., Bays, P. M., Joseph, S., & Husain, M. (2013). Rapid forgetting prevented by
734 retrospective attention cues. *Journal of Experimental Psychology: Human Perception and*
735 *Performance*, 39(5), 1224–1231. <https://doi.org/10.1037/a0030947>
736 Pertzov, Y., Manohar, S., & Husain, M. (2017). Rapid Forgetting Results From Competition
737 Over Time Between Items in Visual Working Memory. *Journal of Experimental*
738 *Psychology: Learning, Memory, and Cognition*, 43(4), 528–536.
739 <https://doi.org/10.1037/xlm0000328>
740 Pivik, R. T., Broughton, R. J., Coppola, R., Davidson, R. J., Fox, N., & Nuwer, M. R. (1993).
741 Guidelines for the recording and quantitative analysis of electroencephalographic activity
742 in research contexts. *Psychophysiology*, 30(6), 547–558.
743 <https://doi.org/10.1111/j.1469-8986.1993.tb02081.x>
744 Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*,
745 32(1), 3–25. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7367577>
746 Posner, Michael I. (1978). *Chronometric explorations of mind*. Oxford, England: Lawrence
747 Erlbaum.
748 Remington, R. W., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by
749 abrupt onsets. *Perception & Psychophysics*, 51(3), 279–290.
750 <https://doi.org/10.3758/BF03212254>
751 Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template
752 estimation for unbiased longitudinal image analysis. *NeuroImage*, 61(4), 1402–1418.
753 <https://doi.org/10.1016/j.neuroimage.2012.02.084>
754 Rosen, A. C., Rao, S. M., Caffarra, P., Scaglioni, A., Bobholz, J. A., Woodley, S. J., ... Binder, J.
755 R. (1999). Neural Basis of Endogenous and Exogenous Spatial Orienting □: A Functional
756 MRI Study. *Journal of Cognitive Neuroscience*, 11(2), 135–152.
757 <https://doi.org/10.1162/089892999563283>
758 Schneider, D., Barth, A., Getzmann, S., & Wascher, E. (2017). On the neural mechanisms
759 underlying the protective function of retroactive cuing against perceptual interference:
760 Evidence by event-related potentials of the EEG. *Biological Psychology*, 124, 47–56.
761 <https://doi.org/10.1016/j.biopsycho.2017.01.006>

762 Schneider, D., Mertes, C., & Wascher, E. (2016). The time course of visuo-spatial working
763 memory updating revealed by a retro-cuing paradigm. *Scientific Reports*, 6, 21442.
764 <https://doi.org/10.1038/srep21442>

765 Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting Attention Within Visual
766 Short-Term Memory: Development and Mechanisms. *Child Development*, 85(2), 578–592.
767 <https://doi.org/10.1111/cdev.12150>

768 Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13
769 years of the retro-cue effect. *Attention, Perception, & Psychophysics*, 78(7), 1839–1860.
770 <https://doi.org/10.3758/s13414-016-1108-5>

771 Spaak, E., Watanabe, K., Funahashi, S., & Stokes, M. G. (2017). Stable and Dynamic Coding
772 for Working Memory in Primate Prefrontal Cortex, 37(27), 6503–6516.
773 <https://doi.org/10.1523/JNEUROSCI.3364-16.2017>

774 Stokes, M. G. (2015). ‘Activity-silent’ working memory in prefrontal cortex: a dynamic coding
775 framework. *Trends in Cognitive Sciences*, 19(7), 394–405.
776 <https://doi.org/10.1016/j.tics.2015.05.004>

777 Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Article
778 Dynamic Coding for Cognitive Control in Prefrontal Cortex. *Neuron*, 78(2), 364–375.
779 <https://doi.org/10.1016/j.neuron.2013.01.039>

780 Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a
781 user-friendly application for MEG/EEG analysis. *Computational Intelligence and*
782 *Neuroscience*, 2011, 1–13. <https://doi.org/10.1155/2011/879716>

783 Tanoue, R. T., & Berryhill, M. E. (2012). The mental wormhole: Internal attention shifts without
784 regard for distance. *Attention, Perception, and Psychophysics*, 74(6), 1199–1215.
785 <https://doi.org/10.3758/s13414-012-0305-0>

786 Taulu, S., Kajola, M., & Simola, J. (2004). Suppression of interference and artifacts by the
787 signal space separation method. *Brain Topography*, 16(4), 269–275.
788 <https://doi.org/10.1023/B:BRAT.0000032864.93890.f9>

789 Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*,
790 123, 77–99. <https://doi.org/10.1016/j.actpsy.2010.02.006>

791 Van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory
792 models. *Psychological Review*, 121(1), 124–149. <https://doi.org/10.1037/a0035234>

793 Van Essen, D. C., Lewis, J. W., Drury, H. A., Hadjikhani, N., Tootell, R. B. H., Bakircioglu, M., &
794 Miller, M. I. (2001). Mapping visual cortex in monkeys and humans using surface-based

795 atlases. *Vision Research*, 41(10), 1359–1378.
796 [https://doi.org/https://doi.org/10.1016/S0042-6989\(01\)00045-1](https://doi.org/https://doi.org/10.1016/S0042-6989(01)00045-1)
797 Vandenberghe, R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2001). Functional
798 Specificity of Superior Parietal Mediation of Spatial Shifting, 673, 661–673.
799 <https://doi.org/10.1006/nimg.2001.0860>
800 Williams, M., Hong, S. W., Kang, M.-S., Carlisle, N. B., & Woodman, G. F. (2013). The benefit
801 of forgetting. *Psychonomic Bulletin & Review*, 20(2), 348–355.
802 <https://doi.org/10.3758/s13423-012-0354-3>
803 Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., ...
804 Brain, F. (2002). Transient neural activity in human parietal cortex during spatial attention
805 shifts. *Nature Neuroscience*, 5(10), 995–1002. <https://doi.org/10.1038/nn921>
806
807

808

809 **METHODS**

810 **Subjects.** Sixty-four (age=22.84 $\pm$ 2.75 years; 21 males), twenty
811 (age=21.07 $\pm$ 2.69 years; 6 males) and twenty-two subjects (age=21.32 $\pm$ 2.50
812 years; 10 males) were recruited in experiments 1, 2 and 4, respectively.
813 Twenty-seven subjects (age=21.43 $\pm$ 2.18 years; 14 males) participated in
814 experiment 3. All had a normal or corrected-to-normal vision and had no history
815 of psychiatric or neurological disorders. Written informed consent was provided
816 by all participants prior to the experiment. Experimental protocols were
817 approved by the local research ethics committee.

818

819 **Stimuli and experimental tasks.** The procedure is shown in Fig.1. Each trial
820 started with a central fixation dot for 500 ms, followed by a memory display for
821 another 500 ms. The display consisted of two or four Gabor patches (radius 5°,
822 contrast 100%, spatial frequency 2 cycles/degree) within an imaginary
823 Cartesian coordinate system. Their orientations were chosen at random
824 differed by at least 10° in each trial. The distance from the center of each Gabor
825 patch to the axes was 6.6°. After a delay period lasting 2,000 ms, a probe with a
826 randomly oriented Gabor patch was presented at the position that was early
827 encoded. Participants were asked to recall the orientation of the mnemonic
828 Gabor patch as precise as possible by rotating the probe using a mouse. In
829 retro-cue trials, an endogenous retro-cue (a white arrow at the center, 1.8° by
830 1.28°) or an exogenous retro-cue (a dashed white circle at the probed location,
831 radius 5°) was presented for 100ms at the middle of the delay period. Following
832 the retro-cue, a second delay period was presented for another 900ms. In no
833 cue trials, the white fixation dot remained on the screen during the whole delay
834 period, without any changes to it. The inter-trial interval was 1,000ms. At the
835 beginning of the experiment, participants were instructed to maintain fixation on
836 the central dot with minimum eye blinks during the first 3000ms in each trial.

837 In experiments 1, 2 and 4, the validity of retro-cues was always 100%
838 correct, meaning that the probe always appeared at the cued position. Subjects
839 were encouraged to make full use of the retro-cue (i.e., only covertly attend to
840 the cued Gabor patch while ignoring the other Gabor patch(es)) in the retro-cue

841 trials. Before the experiments, participants were requested to perform 20
842 practical trials. In both experiment 1 and 2, within-subject factors with cue types
843 (endogenous cue vs. exogenous cue vs. no-cue) and memory loads (load-2 vs.
844 load-4) were randomly mixed within each block. The whole experiment was
845 divided into 8 blocks, each of which contained 72 trials.

846 In experiment 4, memory load was restricted to the load-2 condition, three
847 cue types were randomly mixed within each block, each of which contained 72
848 trials. There were 23 blocks in total.

849 In experiment 3, one of the two retro-cue types (endogenous or exogenous)
850 and its corresponding cue validity (100% correct or 50% correct) were
851 manipulated in different blocks. These four types of blocks alternated for twice,
852 and thus there were 8 blocks in total. Each block contained 80 trials, and
853 four-fifth of them were retro-cue trials, leaving the no-cue condition randomly
854 mixed in the remaining trials. Prior to experiment, subjects performed a brief
855 practice (10 trials per block type) outside of the MEG scanner, following the
856 same order as in the experiment to become familiar with the task. Block
857 sequences were counterbalanced between subjects.

858 Notably, Gabor patches in the memory display were presented at any two
859 of four quadrants under the load-2 condition in experiment 1, while in following
860 three experiments, the two Gabor patches were bilaterally presented at the
861 screen to keep the balance of visual inputs at both sides.

862

863 **Behavioral analysis.** For each condition, the recall error was calculated by
864 subtracting the response angle from the probe angle. Then, the standard
865 deviation (Raw SD) of recall errors was calculated. Both the recall error and
866 Raw SD values were adjusted for circular data by means of the CircStat toolbox
867 for MATLAB® (Berens, 2009). Repeated measures ANOVA on Raw SD were
868 conducted to identify the main effect of cue type (ATTENTION, experiment 1-4),
869 memory load (LOAD, experiment 1-2), cue validity (experiment 3), stimulating
870 position/time point (experiment 4), and their interactions. Adjusted p-values
871 (p_{adj}) were given by FDR correction when applying multiple comparisons.
872 Partial eta squared (η_p^2) and Cohen's d were given as measures of effect size.

873 In experiment 1, six participants were excluded due to poor performance
874 (lower end in the group 99% confidence interval, the same below) in at least
875 one condition or missing data, resulting in 58 subjects for further analysis. In
876 experiment 2, only one subject was excluded from analysis due to poor
877 performance. In experiment 3, Data acquisition for one subject was aborted
878 upon the request of the participant. Another two subjects were excluded from
879 data analysis due to poor performance. Thus, we analyzed behavioral data
880 from 24 participants in experiment 3. In experiment 4, two out of twenty-two
881 subjects were excluded from analysis due to poor performance.

882

883 **Pupil data analysis.** In experiment 1, we used an Eyelink 1000plus eye
884 (SR Research, Canada) to monitor the trajectory of eye movements. Data from
885 each participant's dominant eye was used. The pupil diameter was corrected
886 by the baseline using the mean value of 100ms prior to the memory display on
887 a trial-by-trial basis for each participant. To reduce the difference in pupil
888 changes evoked by the differential sensory format of the cues, we focused on
889 exploring the interaction between memory loads and cue types. Therefore, we
890 conducted a 2-by-2 measures ANOVA with LOAD (load-2 vs. load-4) and
891 ATTENTION (endogenous cue vs. exogenous cue) as independent variables
892 at each time point. In a control analysis to eliminate the contamination from
893 eye movements, good trials were selected off-line, in which fixation was
894 maintained within 1.5° visual angle and no blink was observed throughout the
895 first 3000ms in each trial. Finally, data from 34 participants were qualified and
896 remained for further analysis, with at least 50 acceptable trials remaining in
897 each experimental condition.

898

899 **EEG recording and preprocessing.** In experiment 2, we recorded the brain
900 activities using electroencephalography (EEG). EEG signals were recorded
901 continuously from 32 Ag/AgCl active electrodes (Easycap; Berlin, Germany)
902 according to the international 10/20 System (Pivik et al., 1993). Vertical
903 electrooculogram (VEOG) was measured by an additional electrode applied
904 below the right eye. A BrainAmp DC-amplifier (BrainProducts; Gilching,
905 Germany) sampled EEG and VEOG signals with a frequency of 1000Hz. A

250Hz low-pass filter was used and the impedance of electrodes was kept below 5k Ω during recording.

EEGLAB (Delorme & Makeig, 2004) was mainly used for data analysis. All channels were re-referenced offline to the averaged mastoids by means of the signal recorded from electrodes TP9 and TP10. EEG data were filtered with both 0.5Hz high-pass and 40Hz low-pass filters, and then divided into segments ranging from 700ms before to 3000ms after the onset of the memory display. The mean value of 200ms prior to the memory display was served as the baseline on trial-by-trial basis. Epochs containing artifacts, such as blinks or saccades, and excessive noise ($\pm 75 \mu V$) at any electrode within -200 and +2500ms time window were excluded from further analyses. After this operation, EEG data from 15 subjects were qualified with at least 50 trials in each condition. To obtain contra- and ipsilateral activities evoked by retro-cues, electrodes from both hemispheres were exchanged in trials where subjects were cued to recall the right side of the memory display. Then, the mean amplitudes across trials at each electrode were calculated. As a result, the curves from the right hemisphere stood for contralateral activities, and that from the left hemisphere stood for ipsilateral activities.

ERPs analysis. ERPs were referred to the onset of retro-cues. Time ranges for analyzing PCN and Pd at posterior electrodes P7 and P8 were accordingly set to 350 – 450 ms and 580 – 680 ms, ADAN at anterior electrodes F3 and F4 were tested in the time window of 300 – 500 ms, in line with previous studies (Schneider, Barth, Getzmann, & Wascher, 2017; Schneider, Mertes, & Wascher, 2016). Posterior P1pc and anterior P1ac were identified when amplitudes approached their peaks within time ranges for P1 (100-160ms in our study). For statistical analyses, we conducted 2-by-3-by-2 repeated measures ANOVAs with LOAD (load-2 vs. load-4), ATTENTION (endogenous cue vs. exogenous cue vs. no cue) and hemispheres (contralateral vs. ipsilateral to cued side) as within-subject factors. Lateralized effects were further ensured using paired *t*-tests (i.e. contralateral vs. ipsilateral to the cued side) for each condition. To investigate lateralized differences between two types of retro-cues, paired *t*-tests were calculated for each memory load with

mean amplitudes or latencies of event-related lateralization (ERL) within the corresponding time window as dependent variables.

941

MEG procedure and preprocessing. In experiment 3, anatomical MRI scans were always conducted after MEG recordings. A T1-weighted image was acquired with a Siemens 3-Tesla Prisma scanner (192 sagittal slices, 1mm thick, TR=2.53s, TE=2.98ms, flip angle=7°, FOV=22.4 cm *25.6cm) in 17 subjects or with a GE scanner (192 sagittal slices, 1mm thick, TR =6.64ms, TE=2.93ms, flip angle=12°, FOV=25.6cm *25.6cm) in another 10 subjects (out of 27). The cortex with 1 mm *1 mm *1 mm resolution was then extracted using the FreeSurfer toolbox, for estimating the source activity in the brain (Reuter, Schmansky, Rosas, & Fischl, 2012).

MEG data were obtained with a whole-head 306-channel Vector view system (Elekta-Neuromag, Helsinki, Finland), consisting of 102 magnetometers and 204 orthogonal planar gradiometers. The signal was recorded at a sampling rate of 1000Hz with an online bandpass filter from 0.1 to 250Hz. The head position was measured at the beginning of the experiment with six head position indicator coils. Anatomical landmarks (nasion, brow, left and right ear) and extra points (~100) of the head shape were obtained using a 3D digitizer (Fastrak Polhemus, Colchester, VA, USA). External noise was removed with a signal space separation (SSS) method implemented with MAX filter software (Taulu, Kajola, & Simola, 2004). Head positions in the following runs were re-aligned to the position measured in the first run. During this step, MEG data from two subjects were further excluded due to a technical failure, thus we analyzed the remaining MEG data from 22 subjects. Further analyses were conducted with custom scripts based on Brainstorm toolbox (Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011) for MATLAB®. Continuous data were notch filtered using 50, 100, and 150Hz filters, and then band-pass filtered between 1 and 100Hz. Muscular artifacts and eye blinks were evaluated with ICA, and were visually inspected and manually removed. Remaining data were then cut into 2600ms corresponding to 100ms before the memory display and 2500ms after the onset of memory display. Event-related field (ERF) analysis was conducted on 102 magnetometer channels.

972

973 **Source localization and selection of ROIs.** MEG signals in the empty room,
 974 serving as covariance for noise modeling, were recorded prior to the formal
 975 experiment. We calculated the forward model using the method of overlapping
 976 spheres for each subject, and then applied the Minimum norm model (current
 977 density map) to estimate source activities for trials of each block. Next, neural
 978 activations at the source level for each condition were extracted and then
 979 averaged across trials before performing low-pass filter (below 32Hz) and
 980 z-score transformation. After this, absolute values for each condition from per
 981 subject were calculated and projected onto the MNI template. Two clusters
 982 including the right intraparietal sulcus (IPS) and left DLPFC displayed
 983 significantly larger neural activation in both cued conditions vs. no-cue
 984 condition in the post-cue period of 0-500 ms (permutation test $\alpha < 0.05$). To
 985 compare differences between two types of retro-cue conditions, paired *t*-tests
 986 were used with their latencies of brain responses within specific time windows
 987 (i.e., 100-300ms or 300-500ms after cue onset) as dependent variables.
 988 Pearson correlations between neural activations and behavior were applied to
 989 these brain regions, respectively.

990

991 **Granger causality analysis.** It has been well documented that the superior
 992 parietal lobe (SPL) and inferior parietal lobe (IPL) are involved in voluntary and
 993 involuntary attentional processing, respectively (Corbetta & Shulman, 2002).
 994 To evaluate how the information from visual cortex interacted with these two
 995 regions, and also, how it received regulation from prefrontal cortex relating to
 996 behavior, the lateral occipital cortex (LOC) and dorsal lateral prefrontal cortex
 997 (DLPFC) were recruited for the analysis. To achieve this, we extracted time
 998 courses of these four brain regions, within a post-cue 1000ms time period,
 999 based on the Desikan-Killiany labeling system (Desikan et al., 2006). We then
 1000 put them into the Multivariate Granger Causality (mvgc) toolbox (Barnett &
 1001 Seth, 2014) to perform the Granger causality analysis, where a length of
 1002 200ms per time window stepped by 100ms was applied. Multiple comparisons
 1003 were corrected using Bonferroni correction at a significant level of 0.05.

1004

1005 **TMS procedure.** In experiment 4, anatomical T1-weighted magnetic
1006 resonance imaging (MRI) scans were always conducted preceding TMS
1007 sessions in a different day (interval >48h). They were acquired with a Siemens
1008 3-Tesla Prisma scanner (192 sagittal slices, 1mm thick, TR=2.53s, TE=2.98ms,
1009 flip angle=7°, FOV=25.6 cm *25.6cm) at the ECNU MRI Research Center.
1010 These images were then imported into the BrainSight neuronavigation
1011 software (BrainSight 2.0, Rogue Research, Montreal, Canada) to allow for
1012 stereotaxic registration of the coil with the brain. TMS was delivered via
1013 Magstim Rapid2 stimulator and a 70-mm figure-of-eight coil (The Magstim
1014 Company, Whitland, UK). The TMS intensity for each subject was calculated
1015 as 90% of phosphene threshold (output= 55%±6.5%). The threshold was
1016 determined as subjects perceived phosphenes in 5 out of 10 TMS pulses given
1017 at the lateral occipital (LO) region, localized to be 1–1.5 cm caudal on the skull
1018 in a direct line towards the inion in accordance with various anatomical and
1019 functional maps (Van Essen et al., 2001).

1020 During the experiment, TMS was applied to either the left dorsal lateral
1021 prefrontal cortex (IDL PFC, MNI coordinates: x=-29, y= 33.0, z=28.5) or the
1022 right intraparietal sulcus (rIPS, MNI coordinates: x=24.8, y=-64.2, z=41.7) in 20
1023 blocks, whose MNI coordinates were acquired as the maximal activation
1024 observed in experiment 3. Their stimulation sequence was counterbalanced
1025 between subjects. Importantly, we also recruited vertex as a control area
1026 where behavioral performance was supposed to be unaffected by stimulation.
1027 TMS applied to vertex was always arranged at the 1st,12th (middle) and 23rd
1028 (last) blocks to match subject' states in other blocks which may be influenced
1029 by the practice effect as well as the fatigue effect.

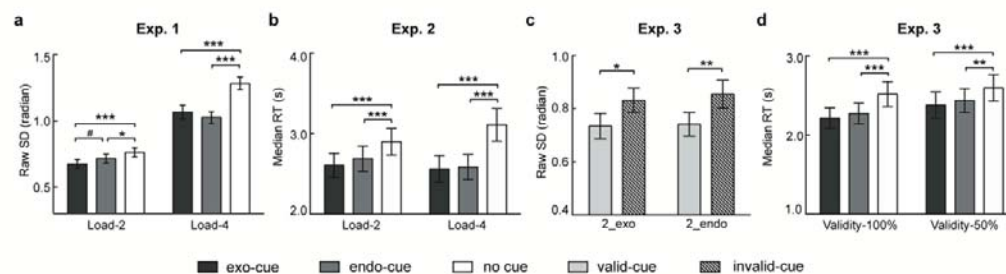
1030 In each trial, a single pulse TMS was applied at one of three time points
1031 (100, 400 and 700ms after the (no) cue onset) in a pseudo-random order. We
1032 collapsed data with different stimulating time points when TMS was applied to
1033 vertex. Therefore, the IDLPFC- or rIPS-targeted stimulation condition (each
1034 cue type * each time point) consisted of 80 trials each, and the vertex-targeted
1035 stimulation condition contained 72 trials per cue type.

1036

1037

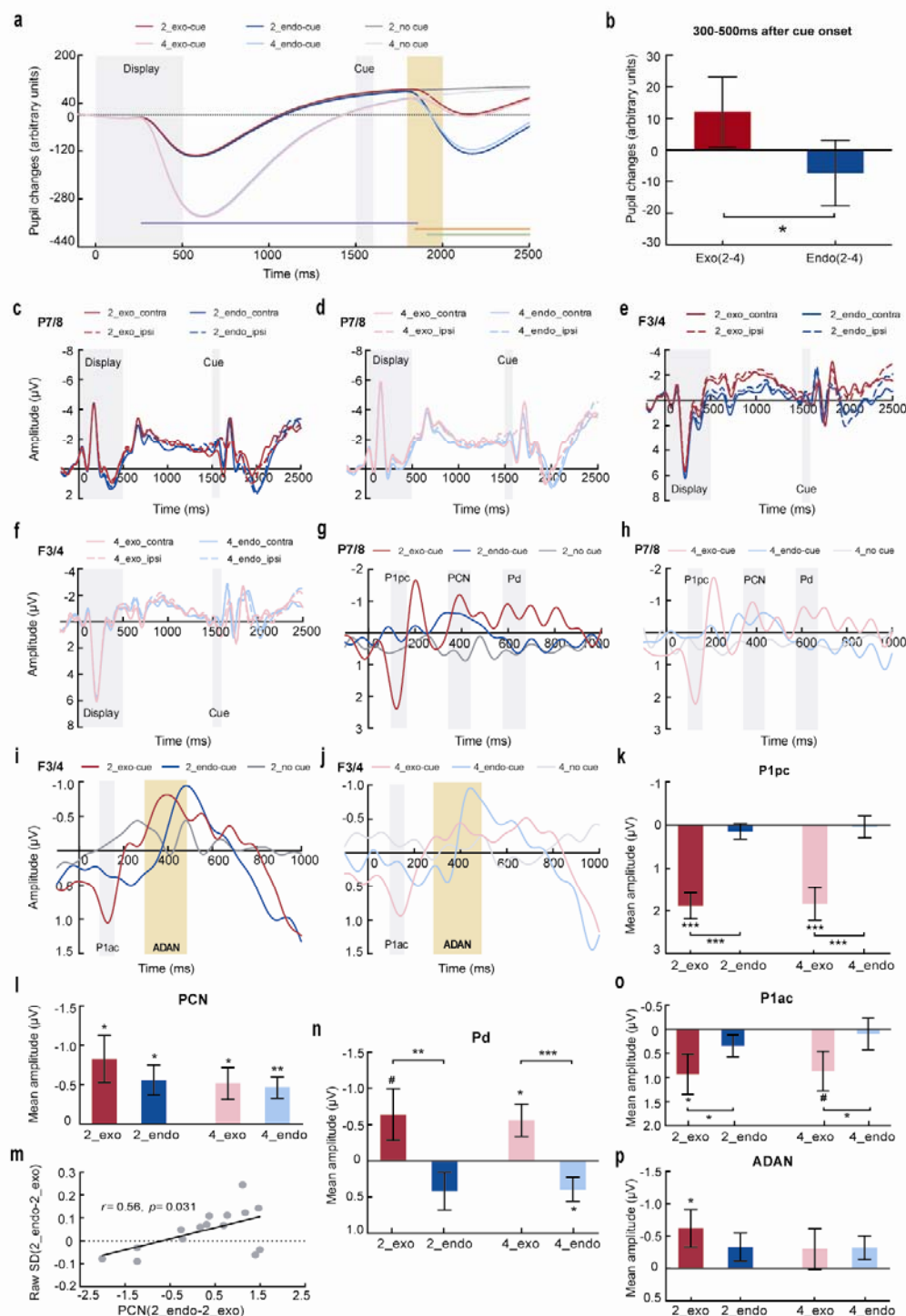
1038

1039 SUPPLEMENTARY FIGURES



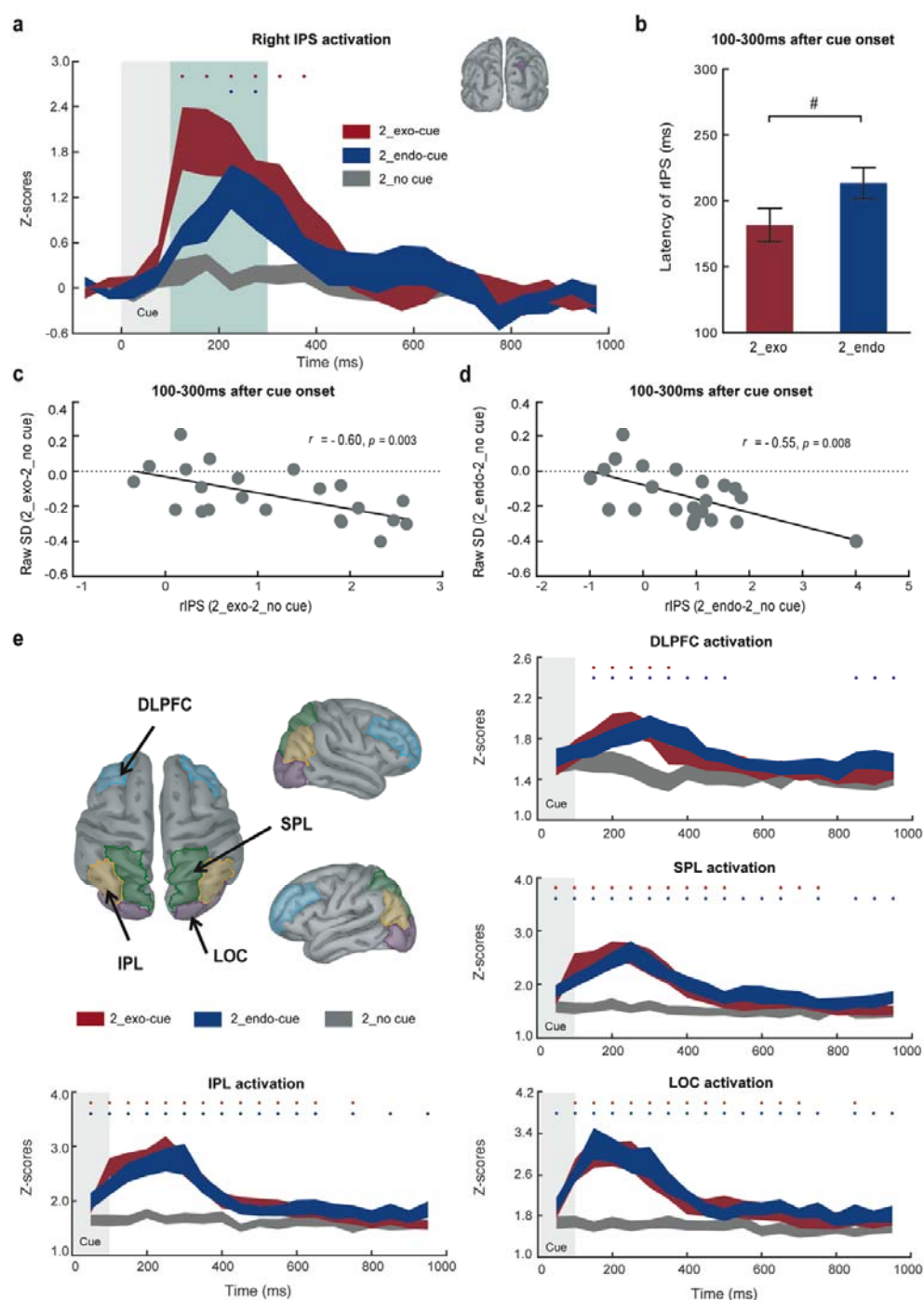
1040

1041 **Supplementary figure 1 | Behavioral results.** **a**, Behavioral results from experiment 1
1042 after controlling eye movements and blinks ($n=34$). **b**, Median reaction times (RTs) from
1043 experiment 2 ($n=19$). **c**, Behavioral results from experiment 3 for retro-cue trials with 50%
1044 cue validity ($n=24$). **d**, Median RTs from experiment 3. # $p<0.1$, * $p<0.05$, ** $p<0.01$, ***
1045 $p<0.001$, all p -values are FDR corrected.



Supplementary figure 2 | Psychophysiological results from experiment 1 and 2. a, Pupil changes using qualified trials from 34 subjects as a function of time referred to the memory display. Purple dots indicate a main effect of LOAD, orange dots indicate a main effect of ATTENTION, and green dots indicate a 2-by-2 interaction. **b,** Load-by-cue interactions within the time window of 300-500 ms after the cue onset. **c-f,** The original

1052 curves referred to memory display at P7/8 electrodes (c-d) and F3/4 electrodes (e-f). **g-j**,
1053 The contra- minus ipsilateral curves time-locked to the cue onset at P7/8 electrodes (g-h)
1054 and F3/4 electrodes (i-j). **k**, Mean amplitudes of P1ac. **l**, Mean amplitudes of PCN. **m**,
1055 Correlations between endo- minus exogenous PCN differences at the load-2 condition and
1056 their behavioral differences at the load. **n-p**, Mean amplitudes of Pd (n), P1ac (o), ADAN
1057 (p). # $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
1058



1059
 1060 **Supplementary figure 3 | Psychophysiological results from experiment 3. a-b,** Grand
 1061 event-related field time-locked to memory display in blocks with 100% validity. Red (blue)
 1062 dots indicate higher brain responses to exogenous (endogenous) retro-cues than to no cue
 1063 condition (one-tailed t -tests, $p < 0.05$). b, The latency of rIPS activation within the range of
 1064 100-300 ms after cue onset for retro-cue conditions. (# $p < 0.1$). **c-d,** Correlations in blocks

1065 with 100% validity between the activation differences of exogenous (c) or endogenous (d)
 1066 retro-cue relative to no cue conditions at rIPS and their behavioral differences within the
 1067 time window of 100-300 ms after the cue onset (two-tailed t -tests, $p < 0.05$). **e**, ROIs on a
 1068 template and z-transferred grand responses within these ROIs in blocks with 100% validity
 1069 time-locked to the cue onset. Red (blue) dots indicate higher responses to an exogenous
 1070 (endogenous) retro-cue than to no cue (one-tailed t -tests, $p < 0.05$).
 1071