

1 Category-based and location-based volitional covert
2 attention affect memory at different timescales

3 Kirsten Ziman^{1,2}, Madeline R. Lee¹, Alejandro R. Martinez¹,

Ethan D. Adner¹, and Jeremy R. Manning^{1,†}

¹Dartmouth College

²Princeton University

[†]Address correspondence to jeremy.r.manning@dartmouth.edu

4 March 25, 2023

5 **Abstract**

6 Our ongoing subjective experiences, and our memories of those experiences, are shaped by our prior
7 experiences, goals, and situational understanding. These factors shape how we allocate our attentional
8 resources over different aspects of our ongoing experiences. These attentional shifts may happen overtly
9 (e.g., when we change where we are looking) or covertly (e.g., without any explicit physical manifestation).
10 Additionally, we may attend to what is happening at a specific spatial location (e.g., because we think
11 something important is happening there) or we may attend to particular features irrespective of their
12 locations (e.g., when we search for a friend's face in a crowd versus a location on a map). We ran a covert
13 attention experiment with two conditions that differed in how long they asked participants to maintain the
14 focus of the categories and locations they were attending. Later, the participants performed a recognition
15 memory task for attended, unattended, and novel stimuli. Participants were able to shift the location of
16 their covert attentional focus more rapidly than they were able to shift their focus of covert attention to
17 stimulus categories, and the effects of location-based attention on memory were longer-lasting than the
18 effects of category-based attention.

19 **Keywords:** covert attention, spatial attention, category-based attention, recognition memory, perception

20 Introduction

21 Our brains' cognitive systems detect and exploit patterns in our prior and ongoing experiences, enabling
22 us to function and adapt in an ever-changing world. However we do not attend to or treat all types of
23 remembered or incoming information equally, and our ability to flexibly adapt our thinking and behaviors
24 can vary markedly with the specific set of concepts or tasks relevant to a given setting or situation (????).
25 There is also substantial variability across people with respect to which aspects of experience (sensory, social,
26 emotional, etc.) are noticed, discriminated between, and acted upon (?). This implies that the same physical
27 (objective) experience may give rise to very different perceived (subjective) experiences across people (??).

28 The aspects of our experience we attend may be under our volitional control or may be unconscious or
29 automatic (?). Both volitional and unconscious attention may be expressed overtly, for example through
30 intentional eye movements (?) or covertly, without any volitional physical change (?). Prior work has
31 explored the similarities and differences in the neural basis of overt versus covert attention (??) as well as
32 the behavioral and neural underpinnings of volitional versus unconscious attention (?) and their differential
33 effects on memory. There is a general consensus that sustained volitional attention enhances memory relative
34 to unconscious attentional processes (??). However, volitional attention takes many forms, such as attention
35 to particular spatial locations or attention to particular visual features or other stimulus properties. How
36 different *types* of volitional attention combine (or compete) to enhance memory remains an open question.
37 Volitional covert attention is of particular interest in that it allows us to dynamically and intentionally
38 manipulate our experience, even when our sensory input remains largely static (i.e., constant physical
39 stimuli, retinal image, etc. ??).

40 Here we examine the ways different types of volitional covert attention interact to affect memory. We
41 designed an experimental paradigm (following ?) that asked participants to attend to a series of presented
42 composite image pairs while keeping their gaze fixed on a central point. The image pairs comprised a left
43 and right image, each constructed by blending an image of a face and place. The stimuli and presentation
44 durations were constant across the two experiments, but the experiments differed in how often we asked
45 participants to change the focus of their attention with respect to image category (face versus place) and
46 image location (left versus right). After the participants attended to a series of images, we used a recognition
47 memory test to assess which aspects of the presented images had been encoded into memory. In both
48 experiments we found that the images participants covertly attended to were better recognized than other
49 images, supporting the notion that attention enhances memory encoding (i.e., they rated attended images as
50 more familiar than unattended images ?). After maintaining the focus of attention to a single image category
51 and location (Sustained Attention Experiment), participants also recognized the attended-category image

52 at the unattended location, and (to a lesser extent) the unattended-category image at the attended location.
53 After more rapidly varying their focus of attention (Variable Attention Experiment), participants showed a
54 similar boost in recognition for the unattended-category image at the attended location, but they did not
55 recognize images at the unattended location. This suggests that participants were able to shift the location
56 of their covert attentional focus more rapidly than they were able to shift their focus of covert attention to
57 stimulus features. We also found differences in the timecourses of these memory effects, suggesting that
58 the impact of location-based attention on memory persists on the order of several seconds longer than the
59 impact of feature-based attention.

60 Materials and methods

61 We ran a total of 53 participants in across two experimental conditions (Fig. ??). The two conditions differed in
62 how often we cued participants to change the focus of their attention. All code and documentation pertaining
63 to our experiments and analyses, along with the experimental stimuli and data, may be downloaded from
64 <http://www.github.com/ContextLab/attention-memory-task>.

65 Participants

66 A total of 53 Dartmouth College undergraduate students enrolled in our study (30 in the sustained attention
67 condition and 23 in the variable attention condition). Following a pilot study using a similar experimental
68 design, we aimed to enroll 30 participants in each condition. However, we fell short of our enrollment
69 target in the variable attention condition when in-person testing was discontinued at our institution due to
70 the COVID-19 pandemic. All participants had self-reported normal or corrected-to-normal vision, memory,
71 and attention. Participants gave written consent to enroll in the study under a protocol approved by the
72 Committee for the Protection of Human Subjects at Dartmouth College.

73 We used a voluntary pre-experimental survey to collect self-reported demographic information about
74 each participant. All 53 participants elected to fill out the survey. Participants ranged in age from 18–21
75 years old (mean: 18.7 years; standard deviation: 0.8 years). Participants reported their genders as female
76 (34 participants), male (18 participants), and gender non-binary (1 participant). Participants reported their
77 ethnicities as Not Hispanic or Latino (44 participants), Hispanic or Latino (7 participants), or declined to
78 respond (2 participant). Participants reported their race as White (37 participants), Asian (13 participants),
79 American Indian or Alaska Native (4 participants), Black or African American (2 participants), and Other
80 (1 participant). Note that each participant could report one or more racial categories, as they deemed

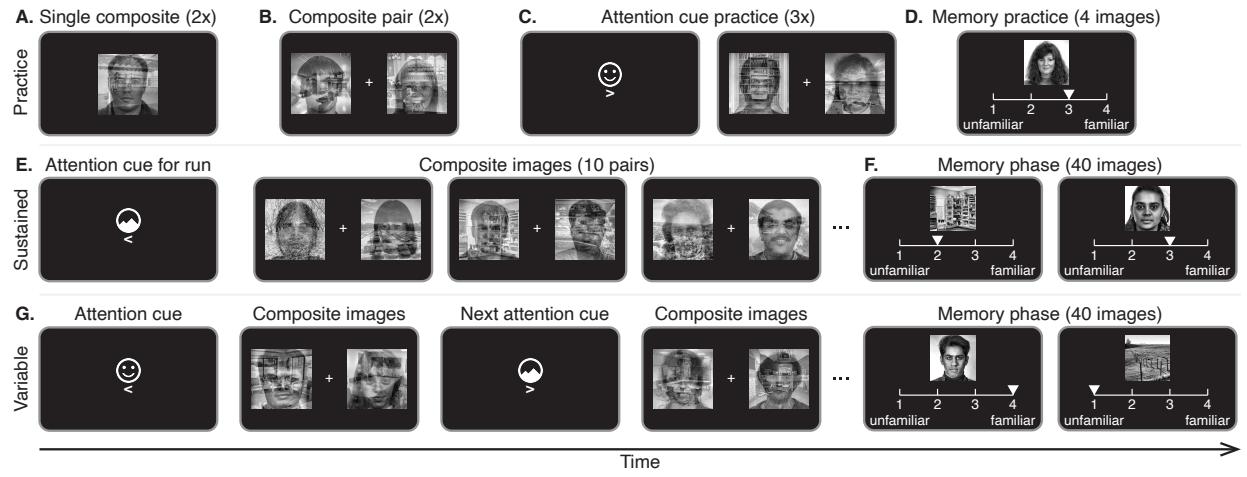


Figure 1. Experimental paradigm. A.-D. Practice phase. A. Composite face/place image. B. A single pair of composite images and a central fixation cross. C. One attention cue practice trial. D. Familiarity judgement practice trial. **E. Sustained attention condition.** Participants receive an attention cue, followed by a sequence of 10 composite image pairs (Panel E), and then they make of familiarity judgements on each of 40 face and place images, presented in sequence (Panel F). **G. Variable attention condition.** Participants study a succession of 10 composite image pairs, each preceded by an attention cue. After studying the images, they make a series of 40 familiarity judgements about face and place images, as in the sustained attention condition (Panel F).

81 appropriate.

82 Fourty-nine participants reported having no reading impairments, and 4 participants reported having
83 reading impairments such as dyslexia. Fifty participants reported having normal color vision and 3 reported
84 having abnormal color vision such as colorblindness. Fifty participants reported taking no medications and
85 having no recent injuries. One participant reported that they had recently “hit [their] head very hard.”
86 Another reported having taken concerta in the past, but mentioned they had not taken it recently. One
87 participant reported using amphetamines regularly, but also clarified that they were not currently on
88 amphetamines at the time of testing.

89 We also asked participants to self-report on their sleep, alertness, and coffee consumption. Participants
90 reported having gotten between 4 and 9 hours of sleep on the night prior to testing (mean: 6.9 hours;
91 standard deviation: 1.3 hours). Participants reported their alertness at the time of testing, and we converted
92 their responses to point values as follows: “Very alert” (5 points), “A little alert” (4 points), “Neutral” (3
93 points), “A little sluggish” (2 points), and “Very sluggish” (1 point). Across all participants, the full range of
94 alertness values were used (maximum: 5; minimum: 1; mean: 3.44; standard deviation: 1.0). Participants
95 reported having consumed between 0 and 2 cups of coffee so far on their testing day (mean: 0.3 cups;
96 standard deviation: 0.5 cups).

97 **Stimulus selection and presentation**

98 Participants viewed photographs of faces, places, and composite images each comprising an equal blend of
99 one face image and one place image. The pool of 360 face images included photographs of adult human male
100 and female faces selected from the FERET database (?). The pool of 360 place images included photographs
101 of indoor and outdoor places selected from the SUN database (?). The images we used from both databases
102 came from a stimulus subset that was manually curated by Megan deBettencourt (personal communication).
103 All images were resized to 256×256 pixels, converted to greyscale, and processed so that every image was
104 matched for mean contrast and intensity. We selected 20 face images and 20 place images from the stimulus
105 pool to use in the instructional and practice phases of the experiments (Fig. ??A–D).

106 In addition to the face and place images, we presented (in white) attention cues to direct the participant’s
107 focus of attention. The attention cues comprised a stylized icon of a face or mountain peaks, directing
108 attention to the face or place component of the images, respectively; and a left- or right-facing angled
109 bracket, directing attention to the left or right image, respectively (e.g., Figs. ??C, E, and G).

110 Our experiment was conducted in a sound- and light-attenuated testing room containing a chair, desk,
111 and 27-inch iMac desktop computer (display resolution: 2048×1152). The participant sat in the chair and

112 rested their chin on a chin rest located 60 cm from the display. The active portion of the display screen
113 occupied 52.96° (width) and 31.28° (height) of the participant’s field of view from the chin rest. Stimuli were
114 sized to occupy 6.7° (width and height) of the participant’s field of view from the chin rest. We maintained
115 a black background (with any text displayed in white) throughout the experiment.

116 **Eyetracking**

117 We recorded participants’ eye gaze positions using a desk-mounted video-based eyetracker with a spatial
118 resolution of 0.1° visual angle root mean squared error and a sampling rate of 30 Hz (Eye Tribe, The Eye
119 Tribe, Copenhagen, Denmark). We calibrated the eyetracker using a 9-point gaze pattern. As described
120 below, we re-calibrated the eyetracker at regular intervals throughout the experiment to protect against
121 camera drift.

122 We used the eyetracking data to home in specifically on behavioral effects related to *covert* attention, as
123 opposed to overt looking effects. Specifically, we excluded from further analysis any images from trials
124 where participants shifted their gaze (for any non-zero amount of time) to any part of the attended composite
125 image during a presentation trial (see Figs. S1, and S2).

126 **Experimental paradigm**

127 Our experiment comprised two testing conditions: a *sustained* attention condition and a *variable* attention
128 condition. Both experimental conditions comprised a practice phase followed by a series of eight task
129 blocks. Each task block was in turn comprised of a presentation phase and a memory phase. The practice
130 and presentation phases differed across the two experiments, and the memory phases were identical across
131 the two experiments. Both experiments were implemented using PsychoPy (?).

132 **Practice phase**

133 Several participants in pilot versions of our experiments reported that they found it difficult to modulate
134 the focus of their attention quickly on command. We therefore designed a practice sequence to orient the
135 participant to the process of quickly modulating the focus of their attention without moving their eyes.
136 The experimenter remained in the testing room throughout the practice phase and answered any questions
137 about the experiment. The practice sequence builds up incrementally to provide a gradual on-ramping for
138 the participant prior to beginning the main experimental tasks that we focused on in our analyses.

139 **Practice shifting the focus of category-based attention to elements of a single composite image.** At the
140 start of the practice phase, we instructed the participant to look at a single composite (face-place blend)
141 image at the center of the screen, and to try to bring the face component of the image into greater focus by
142 attending to it (Fig. ??A). After pressing a button on the keyboard to indicate that they had done so, we
143 displayed a second composite image and instructed the participant to bring the place component of the new
144 composite image into focus. Again, they pressed a button to indicate that they had done so.

145 **Practice shifting the focus of category-based and location-based attention while viewing two composite**
146 **images.** Next, we asked the participant to stare at a fixation cross presented in the center of the screen while
147 two composite images were displayed on the left and right side of the screen, respectively (Fig. ??B). We
148 first instructed the participant to attend to the place component of the left image without moving their eyes.
149 Participants practiced shifting their attention, and they pressed a button on the keyboard to indicate that
150 they had done so. We then displayed a second pair of composite images and instructed the participant to
151 attend to the face component of the right image. Again, the participant shifted their attention in a self-paced
152 manner, and pressed a button to indicate when they had successfully done so.

153 **Practice sustaining category-based and location-based attention over a series of composite image pairs.**
154 We asked participants in the sustained attention condition to practice holding their focus of category-based
155 and location-based attention constant (to the face component of the right image) while viewing a series of
156 three composite image pairs presented in succession (Fig. ??C).

157 **Practice varying category-based and location-based attention over a series of composite image pairs.**
158 We asked participants in the variable attention condition to practice varying their focus of category-based
159 and location-based attention while viewing a series of three composite image pairs, each presented after a
160 different attention cue (Fig. ??C).

161 **Practice reaction time probe.** After practicing modulating their focus of attention to a series of composite
162 image pairs, we introduced a reaction time probe after each image presentation, whereby we presented
163 either an \times or \circ on either the left or the right of the screen (not shown). We asked the participant to press the
164 1 key as quickly as possible when they saw an \times , or the 3 key as quickly as possible when they saw an \circ . We
165 did not impose a time limit on their responses, other than asking participants to respond as quickly as they
166 were able. Participants practiced three trials of modulating their focus of attention to a pair of composite
167 images (3 s), and reacting as quickly as possible to the \times or \circ symbol presented after each composite image

168 pair. The reaction time probe was intended to keep participants continually engaged in modulating the
169 focus of their attention.

170 **Practice recognition memory task.** Finally, we asked the participant to practice reporting familiarity on a
171 recognition memory task (Fig. ??D). We presented a single face or place image at the center of the screen,
172 and asked them to press a button to indicate how “familiar” the image seemed: 1 (very confident they had
173 not seen the image), 2 (somewhat confident they had not seen the image), 3 (somewhat confident they had
174 seen the image), or 4 (very confident that they had seen the image). We instructed the participant to go with
175 their “gut reaction” in the event that they were unsure of how to respond. We allowed the participant up to
176 2 s to provide their response. We gave participants a total of four practice images to rate.

177 After completing the practice phase of the experiment, the participant read the instructions for the task
178 blocks (described next). The experimenter gave participants a chance to ask any remaining questions about
179 the experiment. After answering the participant’s questions, the experimenter calibrated the eyetracker and
180 exited the testing room.

181 **Task blocks**

182 During each task block we asked the participant to modulate their attention while viewing a series of
183 10 composite image pairs (each followed by a reaction time probe), and then we tested the participant’s
184 memory using 40 familiarity judgements. Each participant completed a total of eight task blocks.

185 **Sustained attention condition: presentation phase (Fig. ??E).** Participants viewed an attention cue (1.5 s)
186 instructing them to attend to either the face or place component of either the left or right images in each
187 to-be-viewed composite pair. Next we displayed 10 composite images in succession (each preceded by a
188 fixation cross and proceeded by a reaction time probe). All possible attention cue pairs appeared exactly
189 twice across the eight task blocks.

190 **Variable attention condition: presentation phase (Fig. ??G).** Participants viewed a succession of 10
191 attention cues (1.5 s), each followed by a fixation cross (1 s), composite image (3 s), and a reaction time
192 probe. The attention cues were selected pseudorandomly across trials within each block, with the constraints
193 that no single attention cue pair could appear more than three times across the 10 composite image pairs
194 within a single task block.

195 **Memory phase (Fig. ??F).** After the presentation phase of each task block, we asked the participant to
196 rate the familiarity (on a 1–4 scale, as during the practice phase) of a succession of face and place images.

197 Each image was preceded by a 1 s fixation cross, and participants had up to 2 s to input their rating of each
198 image. Participants made a total of 40 familiarity judgements, about 20 face images and 20 place images. Of
199 these, 20 of the images (10 faces and 10 places) were drawn randomly from the (attended and unattended)
200 composite images that the participant had viewed during the presentation phase. The remaining 20 images
201 (10 faces and 10 places) were novel images that the participant had not encountered during any part of the
202 experiment. At the end of each memory block, the participant was given the opportunity to take a short
203 break. When they were ready to continue with the next task block, they indicated their readiness to the
204 experimenter. The experimenter then entered the testing room, re-ran the eyetracker calibration sequence,
205 and exited the testing room prior to the next task block.

206 Results

207 We ran a volitional covert attention experiment with two conditions; in the sustained attention condition
208 we asked participants to *sustain* the focus of their attention over a succession of 10 stimulus presentations
209 per block whereas in the variable attention condition we asked participants to *vary* their focus of attention
210 with each new stimulus (also for a total of 10 stimulus presentations per block). Each stimulus comprised
211 a pair of composite images (one on the left and one on the right side of the display), where each composite
212 comprised an equal blend of a unique face and a unique place image. We followed the presentation phases
213 of each experimental block with a memory phase, where participants performed a recognition memory
214 task by rating the familiarity of previously experienced and novel face and place images (see *Experimental*
215 *paradigm*, Fig. ??).

216 We first wondered whether (and how) shifts in covert attention might affect participants' ratings during
217 the recognition memory task (Fig. ??). To ensure that our findings were not conflated with where people
218 were physically looking, we excluded from further analysis any images presented during trials where the
219 participant's gaze touched on any part of the attended composite image (see *Eyetracking*, Figs. S1 and S2).
220 For the remaining trials, the participants kept their gaze focused on a fixation cross at the center of the
221 screen while *covertly* shifting the focus of their attention to the cued category component of the composite
222 image on the cued location. In other words, during these remaining trials, participants' physical (external)
223 experiences of the face and place components of every presented composite image remained relatively
224 constant across trials (up to our ability to accurately measure where participants were looking using the
225 eyetracker).

226 Simply by encoding their prior experiences into memory, we reasoned that participants should rate
227 *any* presented images as more familiar than novel images, regardless of whether they were following the

attention cues. We confirmed that this prediction held in both the sustained ($t(29) = 8.856, p < 0.001$) and variable ($t(22) = 5.144, p < 0.001$) conditions. In addition, to the extent that participants were following the attention cues, there *internal* experiences of each image should depend on their internal focus of attention during each image presentation. For example, we expected that the attended-category component of the composite image at the attended location might be better recognized than the other composite image components. Indeed, participants in both experimental conditions rated these “fully attended” images as more familiar than category-matched image components from unattended locations (sustained: $t(29) = 6.893, p < 0.001$; variable: $t(22) = 6.938, p < 0.001$), location-matched images from the unattended category (sustained: $t(29) = 6.710, p < 0.001$; variable: $t(22) = 7.633, p < 0.001$), unattended images that were neither from the attended category nor the attended location (sustained: $t(29) = 8.470, p < 0.001$; variable: $t(22) = 7.256, p < 0.001$), or novel images they had never seen before (sustained: $t(29) = 10.259, p < 0.001$; variable: $t(22) = 7.874, p < 0.001$). We also observed a “boost” in familiarity over novel images for attended *or* unattended images from the cued category (sustained: $t(29) = 10.578, p < 0.001$; variable: $t(22) = 7.354, p < 0.001$) and location (sustained: $t(29) = 8.917, p < 0.001$; variable: $t(22) = 5.726, p < 0.001$).

We also wondered whether the ways participants attended to or remembered the images might depend on image-specific properties like the images’ categories. We repeated the analysis displayed in Figure ?? separately for face and place images (Fig. S3). The same general patterns held for each category, as when we combined across the two stimulus categories as reported above. For example, fully attended face and place images were both rated as more familiar than the category-matched images from the unattended location (sustained: $ts(29) \geq 3.36, ps \leq 0.002$; variable: $ts(22) \geq 6.205, ps \leq 0.001$), attended-location images from the unattended category (sustained: $ts(29) \geq 5.886, ps \leq 0.001$; variable: $ts(29) \geq 4.277, ps \leq 0.001$), images from the unattended category and location (sustained: $ts(29) \geq 6.628, ps \leq 0.001$; variable: $ts(29) \geq 5.624, ps \leq 0.001$), and novel images (sustained: $ts(29) \geq 5.987, ps \leq 0.001$; variable: $ts(29) \geq 5.132, ps \leq 0.001$). Taken together, the above results suggest that what we remember is guided in part by what we attended to, even after accounting for where we look or what specifically we are looking at.

Splitting participants’ responses to face versus place images also revealed that participants often rated attended (and partially attended) place images as more familiar than attention-matched face images (compare dark versus light bars in Fig. S3). We hypothesized that this might be explainable by some property of the relevant cognitive processes or by properties of the stimuli themselves. To help elucidate this distinction, we examined individual exemplars of the face and place images used in our paradigm (Fig. ??A). By design, the face images had consistent head sizes, viewing angles, expressions, and so on. In contrast, the place images varied more substantially across images. For example, some place images depicted human-made structures; others depicted natural scenes; some depicted indoor views; others depicted outdoor views;

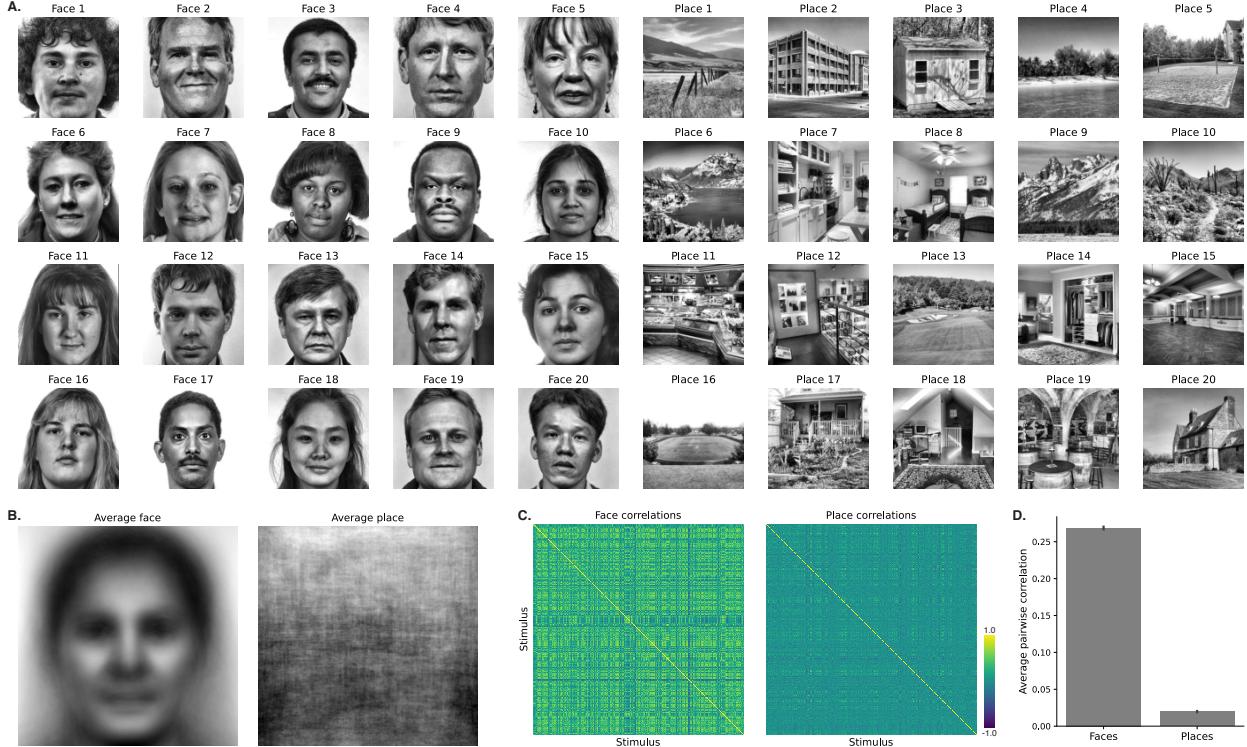


Figure 2. Stimulus examples and properties. A. Example images from each stimulus category. Randomly chosen subsets of 20 face images (left) and 20 place images (right) are displayed. **B. Across-image averages.** Each panel displays the average image, taken across all 360 face images (left) and 360 place images (right). **C. Pairwise correlations.** Each row and column of the matrices displays the correlation (across pixels) in intensity values for one pair of face images (left) or place images (right). **D. Average pairwise correlations.** The bar heights denote the average pairwise correlations between face and place images. Error bars denote across-pair bootstrap-estimated 95% confidence intervals.

etc. This can also be seen by averaging the pixel intensity values across images, separately for the face and place stimuli (Fig. ??B). Whereas the average face image retains many of the landmarks characteristic of most faces (e.g., clearly defined hair, eyes, nose, mouth, head shape, etc.), the average place image does not show place-specific features as clearly, aside from a general tendency for the tops of place images to be lighter than the bottoms of place images. We also computed the pairwise similarities across images from each stimulus category (Fig. ??C) and found that face images tended to be much more similar to each other than place images (Fig. ??D; $t(115258) = 254.764, p < 0.001$). This analysis indicated to us that our experimental paradigm was not well-suited to identifying cognitively meaningful stimulus category differences, since participants' category-specific judgements may be confounded with within-category image similarity differences.

271 Next, we turned to identifying potential differences in participants' behaviors across the two experi-
272 mental conditions. The main difference between the conditions' procedures was in how long participants
273 were asked to maintain the same focus of category-based and location-based attention, across successive
274 image presentations. Therefore, differences in participants' behaviors across the two conditions might
275 reflect differences in the timescales of the relevant cognitive processes. We compared participants' fa-
276 miliarity ratings for images at each attention level across the two conditions. We saw no evidence that
277 people rated fully attended images ($t(51) = -0.649, p = 0.519$), attended-category (but not location) images
278 ($t(51) = 1.163, p = 0.250$), unattended category and location images ($t(51) = -1.600, p = 0.116$), or novel
279 images ($t(51) = -0.435, p = 0.665$) differently across the two conditions. However, participants in the vari-
280 able attention condition rated attended-location (but unattended category) images as more familiar than
281 participants in the sustained attention condition ($t(51) = 2.174, p = 0.034$). This suggests that location-based
282 versus category-based attention may operate over different timescales.

283 **JRM STOPPED HERE**

284 Discussion

285 We ran two covert attention experiments that asked participants to sustain or vary the focus of their covert
286 attention, respectively. When participants held the focus of their feature-based (face versus scene) and
287 location-based (left versus right) attention for a sustained interval, they judged stimuli they had seen as
288 familiar when they overlapped with respect to the features and locations they had attended. The increase
289 in familiarity was larger for attended-feature images than attended-location images. The increase also
290 extended to novel stimuli from the attended image category. By contrast, when participants varied the
291 focus of their feature-based and location-based attention more rapidly, the boost in familiarity for feature-
292 matched stimuli was smaller than that for location-matched stimuli, and did not extend to novel stimuli.
293 Our findings suggest that participants were able to more rapidly modulate their focus of location-based
294 attention than their focus of feature-based attention. The tuning of location-based attention appears to be
295 mediated by enhanced encoding and faster processing at the attended location. The tuning of feature-based
296 attention appears to be mediated by a suppression in the encoding and processing of unattended stimulus
297 features. This suppression effect also affects how new stimuli are processed, and it persists for a longer
298 duration following an interval when the focus of feature-based attention was held constant over a longer
299 duration. Taken together, our findings suggest that feature-based and location-based attention are mediated
300 by different mechanisms and affect memory at different timescales (Fig. ??).

301 The notion that location-based attention operates at a faster timescale than feature-based attention is

Figure 3. Timecourse of encoding and retrieval effects of feature-based and location-based attention.

The focus of location-based attention (green) may be modulated quickly and rapidly enhances memory encoding for nearby stimuli. The focus of feature-based attention (purple) is modulated more slowly, and serves to suppress unattended stimulus features. The effects of location-based attention on memory persist for longer than the effects of feature-based attention. Sustained attention (darker shading) yields more robust enhancement and suppression than shorter-term (variable) attention (light shading). The vertical gray line on the left denotes the duration of a single stimulus presentation (the upper bound by which location-based attention begins to affect encoding, and the lower-bound by which feature-based attention begins to affect encoding). The vertical gray line on the right separates encoding from immediate subsequent retrieval of the encoded information.

302 supported by prior work on the deployment of visual attention (??). Our findings that location-based
303 attention enhances the processing of attended stimuli whereas feature-based attention suppresses the pro-
304 cessing of unattended stimuli is also consistent with prior work on location-based attention (e.g., ?) and
305 feature-based attention (e.g., ?). Our finding that people better remember attended stimuli also follows prior
306 work on interactions between attention and memory (??????). Whereas much of this prior work focused
307 on elucidating the neural basis of these interactions, our work extends these prior studies by elucidating
308 the specific and separable behavioral impacts of feature-based attention (inhibition with a slow onset) and
309 location-based attention (enhancement with a fast onset) on subsequent memory. Both of these effects
310 persisted throughout the 2 min memory phases of both experiments. Therefore future work is needed to
311 elucidate the longevity of these effects beyond 2 minutes.

312 Another important area for future study concerns how the flow of information between different brain
313 structures is modulated according to the focus of volitional attention—particularly with respect to pathways
314 from primary sensory regions (e.g., V1, A1) to regions implicated in encoding ongoing experiences into
315 memory (e.g., medial temporal lobe structures such as the hippocampus and entorhinal cortex, prefrontal
316 cortex, etc.). For example, several studies suggest that attention serves to modulate the *gain* of specific
317 neural circuits (????), effectively facilitating or inhibiting the flow of specific neural representations (??).
318 Prior work suggests that feature-based attention may be supported by changes in connectivity with the
319 thalamus (?), whereas location-based attention may be supported by changes in connectivity with primary
320 visual cortex (?). That feature-based and location-based attention are mediated by different brain structures
321 may explain why these different aspects of attention operate on different timescales and affect memory dif-
322 ferently. A strong test of this hypothesis would entail directly measuring neural activity patterns as people

modulate their focus of attention (e.g., using functional magnetic resonance imaging or electroencephalography), and then using neural decoding approaches (e.g., ???) to follow how neural representations of attended (or unattended) stimuli are transferred from primary sensory regions, to higher order sensory regions, to memory areas. If the effects of attention on memory are mediated by changes in network dynamics, the transmission rates of the representations of attended stimuli from primary sensory regions to memory areas should be facilitated relative to the transmission rates of unattended stimuli. Further, variability in these neural changes (e.g., as a participant focuses their attention with more or less success) should track with behavioral measures of memorability.

Which aspects of our ongoing experiences we choose to attend affects how we process and remember those experiences later. Different forms of attention—e.g., to specific features or spatial locations—operate and affect memory at different timescales, and are likely mediated by different brain networks. Elucidating the behavioral and neural consequences of volitional changes in attention is central to discovering how our thoughts, feelings, goals, and situational understanding fluctuate from moment to moment.

Author Contributions

JRM and KZ developed the concept for this study. Experiment code was written by KZ and ARM, and testing and data collection were conducted by MRL and KZ. KZ, MRL, ARM, and JRM analyzed the data. JRM supervised the project. All authors contributed to writing and editing the manuscript.

Data and code availability

All of the data analyzed in this manuscript, along with all of the code for running our experiment and carrying out the analyses may be found at <http://www.github.com/ContextLab/attention-memory-task>.

Acknowledgements

This work was supported in part by NSF EPSCoR Award Number 1632738 to JRM and by NSF CAREER Award Number 2145172 to JRM. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of our supporting organizations. We are grateful for useful discussions with the EPSCoR Attention Consortium, particularly with Marian Berryhill, Gideon Caplovitz, Patrick Cavanagh, Theresa Desrochers, and Peter Tse. We thank Megan deBettencourt for useful discussions, pointers to our experimental stimuli, and for providing code to facilitate stimulus preparation.

350 We also appreciate useful discussions with Paxton Fitzpatrick, Kevin Hartstein, Talia Manning, Lucy Owen,
351 and Michael Ziman. Finally, we thank Christina Liu and Eowyn Pak for their help running pilot versions
352 of the experiments reported here.