

1 **Learning in the eyes: specific changes in gaze patterns track explicit and**
2 **implicit visual learning**
3

4 **József Arató^{1,2*}, Constantin A. Rothkopf^{3,4} and József Fiser^{1,2}**

5 ¹ *Department of Cognitive Science, Central European University, Budapest Hungary, Nádor utca*
6 ⁹, *Budapest 1051 Hungary*

7 ² *Center for Cognitive Computation, Central European University, Budapest Hungary, Október 6*
8 ^{utca 7}, *Budapest 1051 Hungary*

9 ³ *Centre for Cognitive Science & Institute of Psychology, Alexanderstrasse 10, Technical*
10 ^{University of Darmstadt}, *64283 Darmstadt, Germany*

11 ⁴ *Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1, Goethe University, 60438*
12 ^{Frankfurt}, *Germany*

13 * *Current Affiliation: Cognitive Science Hub, University of Vienna, Liebiggasse 5, Vienna 1010,*
14 ^{Austria}.

15

16

17 Correspondence:

18 József Fiser

19 Department of Cognitive Science

20 Center for Cognitive Computation

21 Central European University

22 Nádor utca 9

23 1051 Budapest

24 Hungary

25 E-mail: fiserj@ceu.edu

26 Tel: +36 1 887 5141

27 **Abstract**

28 What is the link between eye movements and sensory learning? Although some theories
29 have argued for a permanent and automatic interaction between what we know and
30 where we look, which continuously modulates human information- gathering behavior
31 during both implicit and explicit learning, there exist surprisingly little evidence supporting
32 such an ongoing interaction. We used a pure form of implicit learning called visual
33 statistical learning and manipulated the explicitness of the task to explore how learning
34 and eye movements interact. During both implicit exploration and explicit visual learning
35 of unknown composite visual scenes, eye movement patterns systematically changed in
36 accordance with the underlying statistical structure of the scenes. Moreover, the degree
37 of change was directly correlated with the amount of knowledge the observers acquired.
38 Our results provide the first evidence for an ongoing and specific interaction between
39 hitherto accumulated knowledge and eye movements during both implicit and explicit
40 learning.

41

42 **Introduction**

43 Across their lives, people make 2-3 saccades per second during their wake period, which
44 fundamentally determines the sensory information reaching their conscious cognition. Yet,
45 despite an extended literature on the control of eye movements(Findlay & Gilchrist, 2003; Hayhoe
46 & Ballard, 2005; Kowler, 2011; Yarbus, 1967), we have only a rudimentary understanding of how
47 past experiences influence the deployment of attention as indexed by eye movements(Wolfe &
48 Horowitz, 2017). These include observations that gaze biases can emerge from a lifetime of
49 experience, such as taking the inherent uncertainty of the visual system into consideration during
50 visual search(Najemnik & Geisler, 2005), anticipating a ball's trajectory in sports(Brockmole &
51 Henderson, 2006; Land & McLeod, 2000), the tendency to perform visual search from left to
52 right(Spalek & Hammad, 2005), using learnt semantic knowledge(Vö & Wolfe, 2013) or meaning
53 in real world scenes (Henderson et al., 2018). At shorter time-scales, object co-
54 occurrences(Brockmole & Henderson, 2006; Mack & Eckstein, 2011) and episodic memory have
55 been shown to guide visual search(Li et al., 2018). Past experience on an even shorter time-scale
56 can also influence gaze selection, for example when integrating visual information in a given
57 scene with what has been learned about stimulus statistics within minutes(Hoppe & Rothkopf,
58 2016; Yang et al., 2017). A number of these studies investigate jointly how humans develop
59 specific eye movement patterns based on experience with the structure of sensory input and how
60 they use specific eye movement strategies to solve particular tasks(Brockmole & Henderson,
61 2006; Hoppe & Rothkopf, 2016; Land & McLeod, 2000; Li et al., 2018; Mack & Eckstein, 2011;
62 Nelson & Cottrell, 2007; Yang et al., 2017). However, all the above studies considered specific
63 tasks (e.g. categorization, search), and they focused on end results, that is, they showed that
64 after practice, observers learned the identity and/or location of diagnostic features of the task, and
65 their eye movements became more related to these features. Such studies do not clarify, which
66 of the two competing alternatives best describes the nature of the interaction between acquired

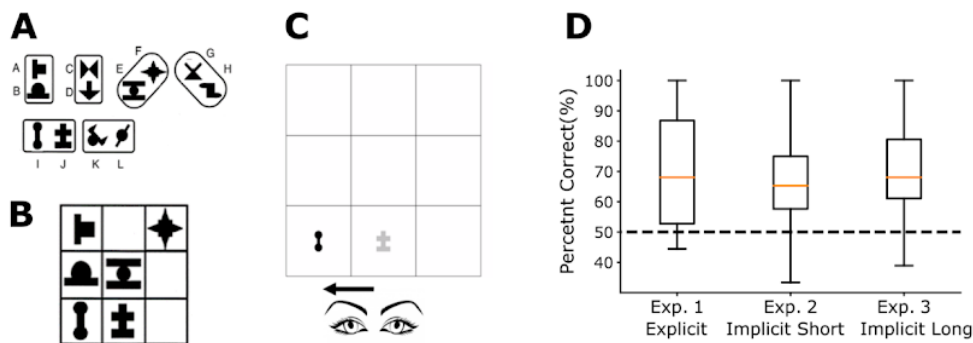
67 knowledge and eye movements in everyday life. First, this interaction could emerge only within
68 the specific context of a clearly defined task via top-down control on sensory processing by high-
69 level explicit knowledge established earlier. Alternatively, the interaction could be a general and
70 ongoing process that modulates human information-gathering behavior all the time and
71 continuously supports both implicit and explicit learning. These two alternatives have very
72 different consequences on how dynamic active perception and the role of eye movements within
73 such perception should be framed with major implications on the relationship between perception
74 and cognition

75 In this study, we address two questions that help evaluate these two alternatives. First, we asked
76 whether there is a difference between how eye movements and sensory learning interact during
77 an explicit task vs. in a task-free observation of structured sensory stimuli. Second, we assessed
78 whether the effect of learning on eye movements is manifested immediately and proportionally
79 with the amount of learned knowledge regardless of this knowledge being implicit or explicit, or
80 alternatively, the effect emerges only after the acquired knowledge becomes explicitly accessible.
81 To explore these issues, we adapted the paradigm of spatial statistical learning (Fiser & Aslin,
82 2001), which allows investigating the process of learning under various levels of implicitness. We
83 altered the paradigm in a gaze-contingent manner, where in each trial, observers saw only a small
84 part of the composite display around their fixation point at a time, and thus through each fixation,
85 they could access different segments of the underlying scene, which consisted of multiple abstract
86 shapes in complex statistical relationships. We coupled this paradigm with either an explicit task
87 (Exp. 1), in which the underlying general structure of the scenes was verbally revealed to the
88 observer prior to the experiment, or under the typical implicit condition of visual statistical learning
89 (Exps. 2&3), where observers had no task other than to explore the unknown scene without any
90 further instructions. This setup allowed investigating, in a continuous manner, the entire process
91 of learning the underlying structure of the scenes from the naive to the expert state, the changes

92 in eye movement patterns during learning, and the effect of explicitness of the knowledge that the
93 observers gathered and relied on.

94 We found that observers' knowledge about the underlying structure of the scenes acquired across
95 multiple presentations induced a specific and significant change in their eye movement patterns.
96 This change reflected the particular spatial structure of the constituents making up the visual
97 scenes, and it progressed proportionally to the amount of learning throughout the learning
98 process. Remarkably, while there was a difference in learning speeds between the conditions,
99 when observers had prior explicit vs. no explicit knowledge, there was no difference between the
100 two conditions in terms of how much a given amount of learning altered the eye movement
101 patterns. Because changes in learning were detectable earlier than changes in gaze patterns,
102 this supports the view that acquired knowledge is integrated continuously into the observer's
103 internal representations without the need for an explicit learning context, and that this knowledge
104 continuously contributes to the control of subsequent information gathering through influencing
105 eye movements.

106



107
108 **Figure 1. Experimental design and test results.** **A)** A set of 12 abstract shapes were randomly assigned
109 to 6 pairs (2-vertical,2-horizontal,2-diagonal) for each participant. **B)** One example of the 144 possible
110 scenes that were assembled from 3 differently oriented pairs randomly arranged on a 3 by 3 grid following
111 the method of previous studies of spatial statistical learning. **C)** Example trial snapshot of the gaze-
112 contingent statistical learning paradigm applied in this paper with the underlying structure of the trial scene
113 shown in B, while the participant's gaze moved from the bottom middle to the bottom left cell (indicated by
114 the arrow). **D)** Results of the 2-IFC familiarity test after the learning phase in the three experiments differing
115 only in instructions and training lengths showed highly significant learning performance (N=40, each, Error
116 bars: full range of data,). Test performance was not different across the three experiments ($F(2,117)=0.89$,
117 $p=.415$, $\eta_p^2=.01$).

118

119 Results

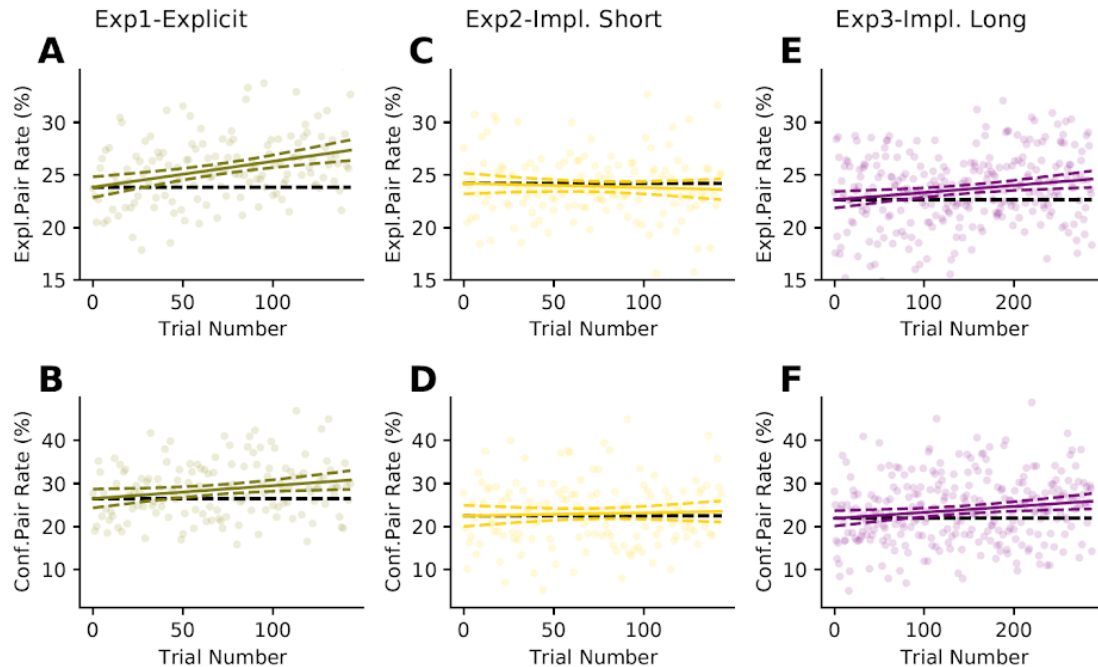
120 **Explicit learning of regularities influences eye-movements.** To establish whether there is an
121 ongoing link between the acquisition of complex environmental regularities and eye-movements
122 during learning, we explicitly revealed the rules of the underlying statistical structure of the
123 presented scenes (but not the identity of shapes in pairs) before Exp. 1. On the 2-IFC familiarity
124 test, participants demonstrated significantly above chance performance (Fig 1D, $M=70.56\%$
125 $95\%CI [64.94, 76.17]$ $t(39) = 7.09$, $p < .001$, Cohen's $d=1.12$), indicating that they, at least partially,
126 acquired the underlying regularities of the training scenes. To investigate the effect of the learned
127 underlying structure on eye-movements, we analyzed whether the exploratory and confirmatory
128 gaze transitions were influenced by the pair structure during training through the slope of
129 regression (β) fitted to the proportion of exploratory and confirmatory looks across trials. The

130 proportion of both types of looks following the pair structures was steadily increasing over the
131 trials (Exploratory: $\beta=.0245$, $p<.001$, Fig 2A; Confirmatory: $\beta=.0301$, $p=.026$, Fig 2B).
132 Furthermore, both measures were predictive of the performance on the final familiarity test on
133 average (Exploratory: $r(38)=0.39$, $p=.013$; Confirmatory $r(38)=0.70$, $p<.001$). Moreover, this
134 predictive power of eye movement patterns on final test performance gradually emerged during
135 the learning phase (Fig 3). To test whether beyond the overall influence, the specific content of
136 learning could also be deciphered from the observer's eye-movements, we used the orientation
137 specific parameters (α_{1-3}) of the the model-based statistical analysis to predict the observer's
138 performance with the differently oriented pairs during the familiarity test. This test showed clear
139 evidence for a significant relationship between the α parameters of eye-movement modulation
140 and learning performance with pairs in all three orientations (Fig 4 A-C). Summarizing the results
141 of Exp. 1, we found that explicit learning of complex regularities can influence eye-movement
142 patterns. Previous evidence on the number of fixations until finding a target (Najemnik & Geisler,
143 2005; Peterson & Kramer, 2001) and looking times (Hoppe & Rothkopf, 2016) suggested that
144 eye-movements can utilize environmental regularities. Our findings extend these results by
145 showing that, with an explicit task, the patterns of explorative eye-movements become sensitive
146 to newly learned spatial stimulus regularities, and the change in eye-movements reflect the
147 amount of learning.

148 **Implicit learning of spatial regularities.** In Experiment 1, we demonstrated a direct link between
149 learning complex regularities and eye-movements when an explicit instruction provided a
150 cognitive support for learning and visual explorations. In Experiments 2 and 3, we investigated
151 whether this link between learning and eye-movements persists when people are solely exposed
152 to the stimuli without any previous knowledge or instructions about regularities within the stimuli.
153 Since learning could only be assessed without interference with implicitness after the end of the
154 exposure period (by the familiarity test), we used two different training lengths in order to assess

155 the link between the strength of learning and its influence on eye-movements at two different
156 stages of learning. Participants demonstrated significant learning in the familiarity test in both
157 experiments (Fig 1D; Exp. 2: $t(39)= 6.81$, $p<.001$, $d=1.08$; Exp. 3: $t(39)= 7.58$, $p<.001$, $d=1.2$),
158 with the performance in Exp. 3 numerically above that in Exp. 2 (Exp. 3 69.65%, [64.64, 74.67]
159 vs.Exp. 2 65.9% [61.38, 70.43]), but this difference was not statistically significant ($t_{78}=1.07$, $p=$
160 $.286$, $d=.24$, Bayes Factor= $.38$).

161 **Change in eye-movements during implicit learning.** Analyzing the effect of the underlying
162 structure on the eye-movements with least-square regression analysis, we found a striking
163 contrast between the two experiments. In Exp. 2, we found no evidence conveyed by regression
164 slopes of any increase in within-pair fixations rates either for exploratory ($\beta=-0.0039$, $p=.513$, Fig
165 2C) or for confirmatory looks ($\beta=0.007$, $p=.643$, Fig 2D). In contrast, and more similarly to Exp. 1,
166 observers' changing fixation rates in Exp. 3 reflected an increasing influence of the pair structure
167 on eye movements over time both in exploratory ($\beta=0.0068$, $p=.005$, Fig 2E) and confirmatory
168 looks ($\beta=0.0139$, $p=.012$, Fig 2F). Compensating the potential confounding effect of variable
169 numbers of eye movements within trials, we reanalyzed the data with a Bayesian mixed model
170 and confirmed the significance of the regression slope in Exp 3, and the lack of such effect in Exp.
171 2.

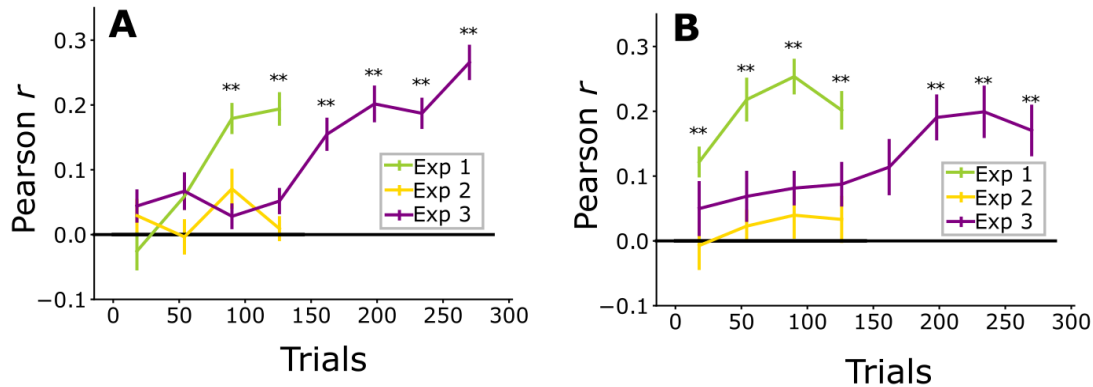


172

173 **Figure 2: Eye-movements are progressively influenced by learned statistical regularities.** Columns
174 indicate the three experiments (Exp. 1: A,B; Exp. 2.: C,D; Exp. 3.: E,F), rows show the two measures
175 (Exploratory and Confirmatory gaze transitions) used to quantify the relation between learned underlying
176 spatial regularities and eye-movement patterns. Dots represent per trial proportion values for each observer
177 for the two measurements, group performance is shown by the least squares regression line (solid) and the
178 95% confidence interval (dashed). Black dashed horizontal line indicates chance performance. **Top Row:**
179 The proportion of explorative eye-movements that were performed according to the statistical structure of
180 the scene (moving from a shape to its pair) was increasing over-time when the instructions were explicit
181 (Exp. 1: **A**, $\beta = 0.0245$, $p < .001$) or during long implicit learning (Exp. 3: **E**, $\beta = 0.0068$, $p = .005$), but it stayed
182 non-significant during the short implicit learning (Exp. 2: **C** $\beta = -0.0039$, $p = .513$). **Bottom Row:** The same
183 conclusions are supported by the Confirmatory Gaze Transitions measure, the proportion of within trial
184 returns to cells already visited on a given trial that were performed within shapes forming pairs. Again, there
185 was a significant increase in Exp. 1 (**B**, $\beta = 0.0301$, $p = .026$, solid line) and Exp. 3 (**F**, $\beta = 0.0139$, $p = .012$),
186 but no change in Exp. 2 (**D**, $\beta = 0.007$, $p = .643$).

187 **Eye-movements predict implicit learning performance.** In Exp. 2, the eye-movement
188 measures were not predictive of the outcome of the familiarity test (Exploratory: $r(38) = 0.17$,
189 $p = .308$; Confirmatory $r(38) = 0.18$, $p = .26$). In contrast, in Exp. 3, both measures had a strong
190 correlation with learning performance (Exploratory: $r(38) = 0.55$, $p < .001$; Confirmatory $r(38) = 0.54$,

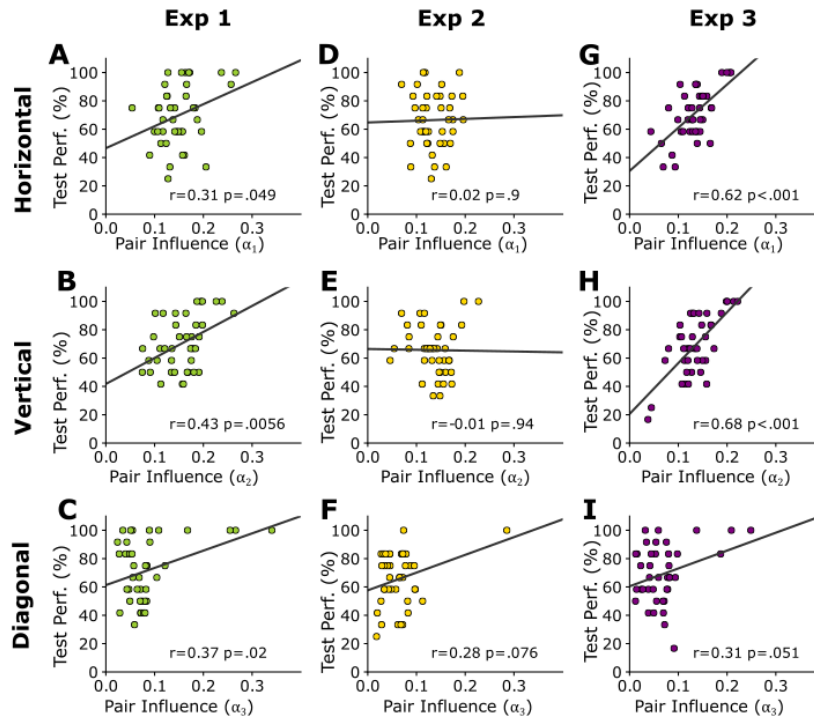
191 $p < .001$). This relationship between learning and eye-movements in Exp. 3 emerged gradually and
192 revealed the strong link only by the second half of the experiment (Fig 3 A-B).



193 **Figure 3: Changes in eye-movements due to acquired knowledge about the statistical structure of**
194 **the stimulus have an increasingly direct link to performance in familiarity tests.** Trial-by-trial eye-
195 movement measures of each participant were correlated with individual learning success measured on the
196 familiarity test. Single trial Pearson r values were averaged in successive 36-trial-long bins. **(A)** Exploratory
197 gaze transitions successfully predicted performance on the familiarity test both in Exp. 1 and Exp. 3.
198 Exploratory looking in all three experiments was not predictive of test performance in the initial bin, but it
199 quickly emerged to a highly predictive level in Exp. 1, unlike in Exp. 2 and in the first half of Exp. 3, where
200 Pearson r values remained at chance. However, in the second half of Exp. 3, a strong relationship between
201 eye-movements and performance emerged matching that of Exp. 1. **(B)** Largely the same pattern of results
202 was found with Confirmatory as with Exploratory transitions, with a faster emergence of statistical influence
203 in Exp. 1, suggesting that returns could reflect a hypothesis testing process of learning. (Error Bars: SEM;
204 ** $p < .01$ after Bonferroni correction).

205 **Eye-movements specifically predict the content of learning.** There was a similar difference
206 between the two experiments in terms of the link between the orientation-specific changes of eye-
207 movements (model α_{1-3}) and familiarity test performance. Predictive relationships were absent in
208 Exp. 2 (Fig 4 D-F), while in Exp. 3, there was a very strong relationship between the magnitude
209 of orientation-specific influence on observer's eye-movements and their pair-specific test
210 performance. For both horizontal and vertical pairs, this effect was strong and highly significant
211 (Fig 4 G-H), while for diagonal pairs, it was weaker and marginally significant (Fig 4 I). We

212 confirmed that these correlations in Exp 3 were not due to general learning effects, but they were
213 highly specific to the particular features the participants learned.



214 **Figure 4. Familiarity test performance is predicted by eye-movement changes due to both implicit**
215 **and explicit learning of stimulus regularities.** On the x axes, parameters of the model-based analysis
216 individually fitted to all gaze-transition data are shown, indicating how strongly a particular pair structure
217 influenced eye-movements relative to the average exploration behavior of the participant. The model had
218 three parameters, corresponding to Horizontal- (α_1 , **Top Row**), Vertical- (α_2 , **Middle Row**), Diagonal-pairs
219 (α_3 **Bottom Row**), representing the relative increase in the number of looks that were in agreement with the
220 spatial arrangement of the pairs. On the y axes, performance on the familiarity test trials containing true
221 pairs from the corresponding orientation is presented. Pearson r and p and Least Square regression line
222 are shown for each condition. The specific link between eye-movements and the content of learning was
223 especially strong in Exp. 3 (Right Column), both for horizontal and vertical pairs. The same two directions
224 also showed a significant relationship in Exp. 1 (Left Column), with a weaker relationship for diagonal pairs
225 due to a stronger ceiling effect. None of the links were significant in Exp. 2 (Middle Column).

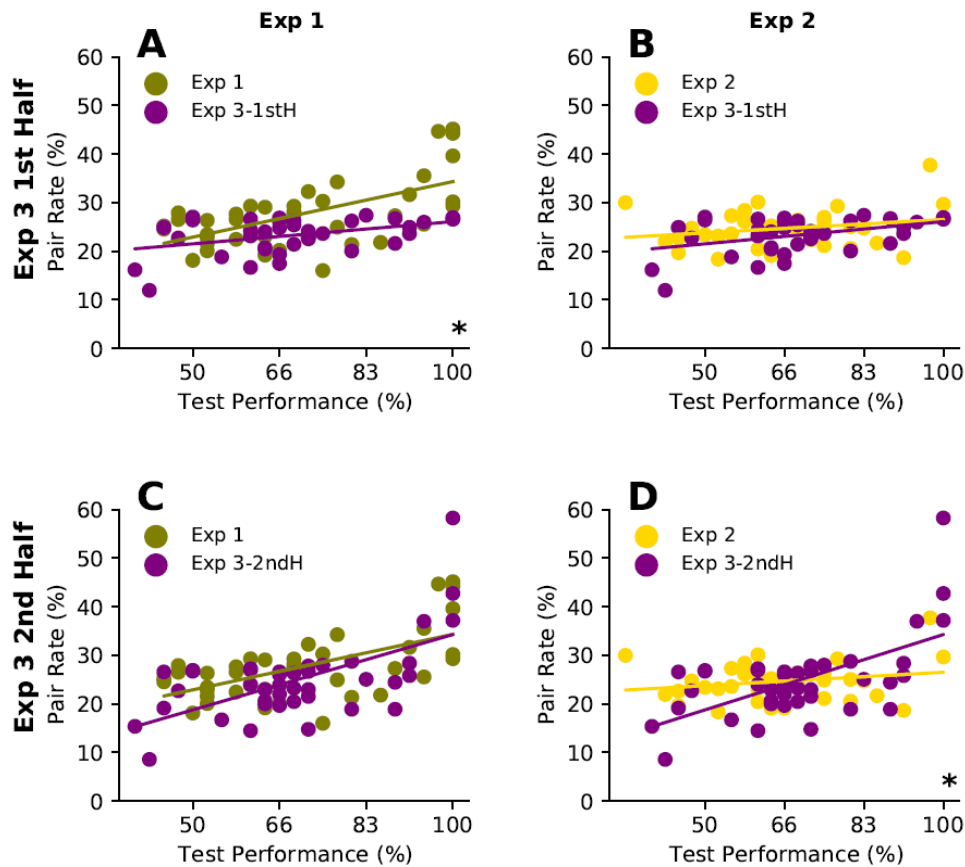
226 **Test similarity of learning influences.** Although our results so far demonstrated that learning
227 both explicitly and implicitly changed eye-movement patterns (Figure 2), it is unclear if these
228 changes in eye-movement were linked only to the amount of learning regardless of whether this

229 knowledge was acquired in an explicit or implicit manner. We hypothesized that, while eye-
230 movements in Experiments 2 and the first half of Experiment 3 were obviously similar, in the
231 second half of Experiment 3, when participants already gained some implicit knowledge of the
232 structure of the input comparable to the gain from explicit instructions in Experiment 1, the eye-
233 movement pattern changes would be indistinguishable for those in Experiment 1. To test this
234 hypothesis, we performed four analyses of covariance (ANCOVA) comparing within-pair eye-
235 movements between Exps 1 or 2 and the two halves of Exp 3, while controlling for the amount of
236 learning. In these analyses, the *Average rate* of within-pair eye-movements of each participant
237 combined across exploratory and confirmatory looks was the dependent variable. The *Type of*
238 *the experiment* was the independent categorical variable, and *Test performance* indicating the
239 amount of learning was the covariate with an interaction term between the covariate and the
240 independent variable.

241 **The relationship of learning & eye-movements across experiments.** Confirming the results
242 in the sections above, the comparisons between Exps 1 and 3 showed that the “Test performance”
243 covariate had a very strong influence on eye-movements (Exp1/Exp3-First half: $F(1,76)=29.22$,
244 $p<.001$, $\eta_p^2=.28$; Exp1/Exp3-Second Half: $F(1,76)=45.67$, $p<.001$, $\eta_p^2=.38$). Comparing the first
245 half of Exp 3 and Exp 1 (**Fig 5A**), we found that this influence of test performance had a significant
246 interaction with the Type of experiment ($F(1,76)=5.36$, $p=.023$, $\eta_p^2=.07$), which rendered the lack
247 of overall main effect of Type of experiment ($F(1,76)=1.64$, $p=.204$, $\eta_p^2=.02$) uninterpretable. By
248 the second half of Exp. 3 (**Fig5C**), neither the slope $F(1,76)=1.05$, $p=.31$, $\eta_p^2=.01$), nor the overall
249 eye-movements were dependent on the Type of the experiment ($F(1,76)=2.04$ $p=.158$, $\eta_p^2=.03$).
250 Thus, gaze-patterns were strongly influenced by the learned knowledge and became
251 indistinguishable between the explicit and implicit experimental conditions. The same analysis for
252 Exp. 3 vs. Exp. 2 showed the opposite pattern. When comparing Exp 2 to the first half of Exp 3
253 (**Fig5B**), the Type of experiment had neither a significant main effect ($F(1,76)=1.33$, $p=.253$,

254 $\eta_p^2=.02$) nor an interaction ($F(1,76)=0.51, p=.477, \eta_p^2=.01$) with the covariate Test performance
255 confirming high similarity across the two conditions. In contrast, when comparing Exp 2 to the
256 second half of Exp. 3 (**Fig5D**), although the Type of experiment had a significant main effect
257 ($F(1,76)= 10.67, p=.002, \eta_p^2=.12$), it also had a significant interaction $F(1,76)= 10.58, p=.002,$
258 $\eta_p^2=.12$) with Test performance indicating that very different causes shaped the gaze-patterns in
259 the two experiments. Importantly, the influence of the Test performance covariate was significant
260 already when the first half of Exp 3 was compared to Exp. 2 ($F(1,76)=8.18, p=.005, \eta_p^2=.1$), but
261 as expected, it became stronger when the second half of Exp 3 was considered ($F(1,76)=21.66,$
262 $p<.001, \eta_p^2=.22$). These analyses indicate that while initially the (relatively weak) relationship
263 between eye-movements and the learning of the underlying structure was very similar between
264 the first half of Exp. 3 and Exp 2, as implicit knowledge accumulated further in Exp 3, it started to
265 influence eye-movements more strongly, and the eye-movement patterns in Exp 3 were
266 influenced in the same way as in the completely explicit learning context of Exp. 1. Thus, these
267 results confirm our hypothesis that in our experiments, the amount of the acquired knowledge is
268 the main driving force behind the changes in eye-movements patterns regardless of the explicit
269 or implicit nature of the experimental conditions.
270

271



272 **Figure 5. The relationship between learning and eye-movements during implicit and explicit learning**
273 **becomes very similar over time.** Scatter plots between familiarity test performance (x-axis) and the ratio
274 of within-pair eye-movements (y-axis) across the three different experiments with Exp 3 splitted to two
275 halves. *Top Row:* Comparison of the first half of Exp 3 to Exp 1 (A) and Exp 2 (B). *Bottom Row:* Comparison
276 of the second half of Exp 3 to Exp 1 (C) and Exp 2 (D). Dots represent mean pair rate combined across
277 Exploratory and Confirmatory looks and the corresponding test performance of individual participants. The
278 lines indicate results of multiple linear regression corresponding to the ANCOVA described in the main text.
279 Stars (*) in the bottom right corners mark a significant interaction between test performance and the
280 covariate Experiment type (A,D), the main focus of this analysis. The main effect of test-performance was
281 significant in all four analyses.

282 **Statistical influence on eye-movements is automatic, but only emerges after sufficient**
283 **learning.** In Exp. 3, we found that, given sufficiently long exposure, learning regularities implicitly
284 and learning them by explicit instructions (Exp. 1) influence visual exploration very similarly.
285 Importantly, the relationship between acquired knowledge and gaze patterns was

286 indistinguishable between the explicit and long implicit conditions. This suggests that the influence
287 of learned knowledge of environmental statistics on eye-movements is automatic, and it does not
288 require a well-defined task or cognitive awareness to emerge. We also found that this effect was
289 tightly linked to the specific knowledge acquired about the statistics of the input. Meanwhile, in
290 the shorter implicit experiment (Exp. 2), we found comparably large learning in the familiarity test
291 without any detectable influence of this learning on eye movements. This provides evidence about
292 the complex relationship between learning and eye movements indicating that precise
293 assessment of the acquired knowledge and good sensitivity in measuring changes in eye-
294 movement patterns will be needed for an ultimate characterization of their relation.

295 **Discussion**

296 Using a novel gaze-contingent statistical learning paradigm, we clarified three aspects of how
297 sensory learning and eye movement patterns interact. First, we confirmed that acquiring
298 knowledge about the underlying structure of the visual environment accumulated by sensory
299 learning can have an effect on the patterns of eye movements even on the short run. Second,
300 we showed that this effect is highly specific to the statistical composition of the incoming sensory
301 input as the knowledge acquired by learning and characterized by different orientations of the
302 underlying chunks could be reliably identified by the individual looking patterns. Finally, we found
303 that, apart from the learning speed, the effect of knowledge on eye movements was independent
304 of whether the observers gained it using explicit instruction about the underlying structure of the
305 input or they obtained it by exploring the scenes without any specific prior information.

306 Previous studies investigating the relationship between environmental regularities and eye-
307 movements fall roughly into two groups investigating complementary aspects of the phenomenon.
308 Studies in the first group focused on the interaction between explicitly or implicitly defined but
309 already available internal knowledge and eye movements in various tasks by investigating the

310 number and position of fixations necessary for finding a target in a display or making a
311 decision(Chukoskie et al., 2013; Hoppe & Rothkopf, 2019; Morvan & Maloney, 2012; Najemnik &
312 Geisler, 2005; Peterson & Kramer, 2001; Yang et al., 2017). Studies in the other group
313 investigated the effect of learning on eye movements but only in terms of learning temporal
314 regularities and adjusting the timing of fixations accordingly(Glimcher, 2003; Hoppe & Rothkopf,
315 2016). Our study is the first to combine these two aspects by investigating the ongoing process
316 of developing an internal representation of the input's spatial structure and showing how the
317 momentary result of this learning process continuously interacts with the pattern of eye
318 movements.

319 Our design also allowed addressing directly the controversial issue regarding the role of explicit
320 vs. implicit knowledge in controlling eye movements. Some studies found that only explicit
321 memories have an influence on eye-movements(Hannula et al., 2012; Smith et al., 2006), while
322 others reported that eye-movements can be used to detect memory traces that are not yet
323 amenable to conscious report(Hannula et al., 2012; Hannula & Ranganath, 2009). Although our
324 findings do not decisively resolve this controversy, we show in a unified setup how eye-
325 movements can reflect memory traces as an outcome of both an explicit and a sufficiently long
326 implicit learning process. While there is an ongoing semantic debate about the definition of explicit
327 vs. implicit memory(Batterink et al., 2015; Greenwald & Banaji, 2017; Roediger, 1990), our results
328 provide two important observations pertinent to the issue. First, the majority of the implicit
329 observers even in the long implicit experiment (Exp 3) did not have explicit access to the gained
330 knowledge about the underlying scene, as indicated by their verbal post-test report. Nevertheless,
331 they showed a monotonic increase in correlation between their implicitly-acquired knowledge and
332 effects on their eye movements. Moreover, even after removing subjects, who performed
333 perfectly on the familiarity test in Exp 3, in a control measure, our conclusion remained the same.
334 Second, the nature of changes in the eye movements in the implicit and explicit conditions were

335 very similar as measured by the rate of exploratory and confirmatory gaze switches, suggesting
336 that the underlying processes were also shared across the two conditions. These two
337 observations indicate that the phenomenon we uncovered is, indeed, a general and automatic
338 process that is driven by knowledge regardless of whether this knowledge is acquired implicitly
339 or explicitly. Moreover, this automatic process influences information collection during perception
340 continuously and in proportion to the amount of acquired knowledge.

341 While the correlational nature of our findings does not allow the assessment of the causal
342 relationship between eye-movements and learning, our results have implications for the suitable
343 framework for capturing the interplay between learning new information and selective data
344 acquisition due to eye movements constrained by internal knowledge. The continuous ongoing
345 nature of the emerging knowledge-based effect on eye movements and the independence of this
346 knowledge of explicitness does not support frameworks positing that eye-movements are affected
347 in an attention-like manner only when the underlying structure of the environment has been
348 learned and it is explicitly accessible. Based on our results, it is more parsimonious to assume an
349 ongoing bi-directional relationship, in which learning influences eye-movements and eye-
350 movements scaffold learning, reflecting a continuous intertwined link between new sensory input
351 and top-down memory related control(Chun & Turk-Browne, 2007; Gottlieb, 2012).
352 Computationally, this process is better represented by dynamically evolving hierarchical inference
353 making(Lake et al., 2015), in which prior knowledge and momentarily collected information is
354 jointly handled for continuously interpreting and controlling sensory input than by two-stage
355 schemes with an initial sweep of bottom-up process followed by specific top-down cognitive
356 filtering(Itti & Baldi, 2009; Schütz et al., 2012).

357 Finally, the method we used in this study can also improve our understanding of the computations
358 involved in statistical learning. Despite being considered as a fundamental form of human
359 knowledge gathering, statistical learning is still not well understood at the process level. This is

360 due to the fact that the majority of studies use the early methodology established in classical
361 papers (Fiser & Aslin, 2001; Saffran et al., 1999), in which the effect of learning is measured on
362 a separate test phase following exposure and this provides only limited information about
363 characteristics of learning (Siegelman et al., 2017). Although recently, different methods were
364 proposed to deal with this problem by tracking the ongoing processes of learning visual
365 regularities (Karuza et al., 2014; Siegelman et al., 2017, 2018), these methods are restricted to
366 temporal statistical learning and raise new concerns due to using explicit instructions instead of
367 truly implicit learning, and increased number of test trials that could interfere with learning. In
368 contrast, our method can be used to track the learning of complex spatial regularities in a natural
369 manner as in the classical experiments, since it relies on an independent and unconscious
370 behavioral measure -eye movement patterns- that does not require changing the original setup
371 of statistical learning and still provides information continuously about the characteristics of the
372 emerging representation.

373 In conclusion, we provided evidence for the first time for a continuous and tight link between
374 human visual information sampling strategies manifested by eye movements and the emerging
375 internal knowledge of environmental regularities. Our results frame natural vision as a process, in
376 which active selection from the incoming information and internal knowledge jointly determine
377 both the interpretation of the input and further changes in internal knowledge.

378

379

380 **Methods**

381 **Participants**

382 Altogether 120 participants naïve about the purpose of the study and about statistical learning
383 were recruited via a local student organization and received monetary compensation for their
384 participation. 40 participants were assigned to each of the three experiments (Exp. 1: age: 25.5
385 +/- 4.6 years, 13 male; Exp. 2: age: 22.1 +/- 2.8 years, 13 male; Exp. 3: age: 23 +/- 5.5 years, 10
386 male). We chose a sample size larger than most previous statistical learning studies (Batterink et
387 al., 2015; Fiser & Aslin, 2001; Turk-Browne et al., 2005) based on power analysis, as we wanted
388 to assess the variability in the individual learning performances. One additional participant
389 completed Exp. 2 but was excluded from the final sample, because upon completing the study
390 revealed not being naïve about visual statistical learning.

391 **Procedure**

392 In Experiment 1, after calibration and practice, but before the start of the main experiment,
393 participants were instructed to explore the scenes and find pairs of shapes that always appear
394 next to each other in a horizontal, vertical or diagonal arrangement. They were also told that they
395 would be questioned about the identity of the pairs afterwards (Explicit instructions). Participants
396 had 6 seconds to explore each of the 144 scenes, presented in a random order, resulting in a
397 total training time of approximately 16 minutes.

398 All aspects of Experiments 2 and 3 were identical to those in Experiment 1 except for the lack of
399 explicit instructions. After calibration and practice, but before the start of the main experiment,
400 participants were told to explore the scenes and pay attention to what they see. They were also
401 told that they will be tested on what they had seen after the exploration phase, however, they

402 were not told about any potential regularity or structure in the stimuli nor about the nature of the
403 subsequent test. These are the canonical conditions of implicit visual statistical learning used in
404 previous studies(Fiser & Aslin, 2001; Turk-Browne et al., 2005). Exp. 2 was the same length as
405 Exp. 1 (~16 mins), but in Exp. 3, the learning phase was double in length: each one of the 144
406 unique scenes were presented once in each half of the experiment in a different random order. In
407 Exp. 3, completing the learning phase took approximately 32 mins, with a short break in the
408 middle, where participants were kindly asked to continue paying attention.

409 All experiments were conducted in a dimly lit and sound attenuated room. A Tobii EyeX 60Hz
410 eye-tracker was calibrated using a seven-point calibration from a viewing distance of 60 cm. After
411 calibration, participants completed ten 6-second-long practice trials, where randomly selected
412 images of dogs were revealed in a gaze-contingent manner within the 3 x 3 grid: the content of
413 each cell was visible only when the observer's gaze fell within the central 5.7 x 5.7 degrees of the
414 cell in two subsequent eye position samples (taken approx. 15 ms apart), otherwise the given cell
415 was shown empty. The trials in the learning phase of each experiment were also 6-second-long
416 and they followed the same gaze contingent rule as during practice.

417 Each trial started by a fixation cross appearing in one of the empty grid cells, where the observer
418 had to fixate to initiate the trial. The position of the fixation cross was uniformly distributed across
419 trials, appearing at the center of each cell of the 3 x 3 grid an equal number of times during the
420 experiment in a random order. Unlike previous spatial statistical learning studies, the full scenes
421 in these trials were never visible at once. Instead, individual shapes were revealed in a gaze-
422 contingent manner, when the participants' gaze was inside the mid-region of a cell. When
423 participants looked at a cell containing the shape, the shape appeared at full contrast as long as
424 the participant's gaze was in the given cell, but gradually faded away becoming invisible within
425 1.5 sec when the participant looked away to a different cell. This way, maximally two shapes of
426 the scene were displayed at any given time and only one of them at full contrast. If the observer's

427 gaze was in the mid-region of a cell not containing a shape in a given trial, a gray rectangle was
428 revealed indicating that the cell was empty in order to reduce the observer's uncertainty whether
429 s/he managed to fixate on the cell. These gray rectangles remained visible until the trial was over,
430 thereby ensuring that the end of each trial was easily noticeable. Participants were free to visit or
431 revisit with their gaze any of the cells during the trial. When the trial was over after 6 seconds, all
432 shapes and gray rectangles disappeared, and after a 500ms inter-trial-interval, the next fixation-
433 cross appeared at one of the cells to initiate the start of the next trial.

434 At the end of the learning phase, after a short break, a two-interval-forced-choice (2-IFC) test
435 session followed, with trials in which participants were told to select the more familiar of the two
436 pair combinations presented based on what they had seen during the learning phase. For the
437 test, 6 foil pairs (with two shapes that never appear in the presented arrangement during learning)
438 were created from the original shapes and those were tested in a fully counterbalanced manner
439 against each of the real pairs of the inventory, resulting in 36 test trials presented in a random
440 order. The within-test trial order of the real versus foil pair was pseudo-randomly balanced across
441 the test. On each trial, participants used the left and right arrow keys for the 1st and 2nd pair,
442 respectively, to indicate which pair was more familiar.

443 **Data Analysis & Measures**

444 All data were analyzed in Python, and statistics were calculated using the SciPy, *scikit-learn*,
445 *Pingouin* and *statsmodels* libraries. Bayes factors were calculated using the method proposed
446 by Rouder et al (Rouder et al., 2009) with an uninformative prior. Since the exact gaze position
447 within the central region of each cell had no functional consequence, eye-movement data was
448 analyzed based on whether or not the fixation samples were within the gaze contingent central
449 region of one of the cells. On average, participants made more than seven (7.2 +/- 1) transitions
450 between the central regions of different cells in a trial. From these transition events, we focused

451 on the ones that were potentially related to learning by using a method detailed below. Since the
452 number of transitions could also change as the learning session progressed, we focused on
453 proportions and not on the absolute number of events.

454 Eye-movement transition data was separated into two different measures that could indicate
455 different behaviors: *exploratory transitions* and *confirmatory returns*. An exploratory transition was
456 defined as a gaze transition to a cell for the first time during a trial, while a confirmatory return
457 was defined as transition to a cell that had already been visited on the current trial. The difference
458 between these events is important, since in case of a return, the participant could be more certain
459 what s/he would see at a given location, as s/he had already seen the content within the last few
460 seconds. In case of an exploratory transition, no such information was available, therefore, the
461 content of the cell could be predicted/expected only if 1) the cell contained a member of a shape
462 pair whose other member the participant already saw during the current trial, AND 2) only if the
463 participant had already learned about the spatial relationships between shapes during the
464 previous trials. Within the exploratory and confirmatory measures, we calculated the proportion
465 of looks that were performed from a shape to its pair, and used this calculation for the assessment
466 of whether the underlying statistical structure had an effect on the transitions. Finally, as a
467 combined measure, we used the rate of within pair eye-movements, which was defined as the
468 proportion of gaze transitions when participants looked from a shape to the cell containing its pair
469 as opposed to other cells.

470 For the analysis of temporal changes in the gaze data across trials, we used regression to predict
471 the eye-movement data with trial number as a predictor. We analyzed the results with two different
472 regression methods, and found support with both of them for the same conclusions. The first
473 method was a simple linear regression predicting the average eye-movement measures across
474 participants (Fig 2). The second was a linear mixed model, predicting a slope for eye-movements
475 across participants, but including a random intercept for each observer.

476 To analyze whether temporal changes in looking behavior across trials were linked to learning
477 (Fig 3), we calculated the Pearson correlation between our eye-movement measures on each trial
478 and the performance in the final familiarity test. Next, we divided the obtained r values in 36 trial-
479 long consecutive bins (yielding 4 bins in Exps. 1 & 2 and 8 bins in Exp. 3), and analyzed whether
480 the r values in each bin were different from zero using a standard one sample t-test. For statistical
481 correction of multiple comparisons, the Bonferroni method was used.

482

483 **Computational Analysis**

484 Our goal was to quantify how much participants' gaze trajectories changed from random
485 exploration to a pattern determined by statistical regularities over the duration of the experiment.
486 We used a model-based analysis to obtain a measure that could be fitted to all gaze transitions
487 without relying on the selection of particular events. For each participant, the model measured
488 the increase of alignment between looking behavior and the statistical structure of the stimuli
489 compared to the average behavior as quantified by the distribution of transition probability across
490 the cells of the grid. Since there were three types of regularities in the stimuli (link across
491 horizontal, vertical, and diagonal orientations), the model had three parameters (α_{1-3}),
492 representing increased gaze transitions between shapes forming pairs in each of the three
493 orientations. For example, the value of α_1 represented an increased probability of looking from
494 shape1, which was a member of a horizontal pair, to the position of shape2, the other shape in
495 the pair. For each observer, the values of the three parameters were fitted trial-by-trial using the
496 maximum likelihood method. To test whether these orientation-specific changes in eye movement
497 behavior during the learning phase could predict performance in the test session, we separated
498 the 36 test trials based on the orientation of the true pair in the trial, yielding 12 test trials for each

499 orientation. Next, we used Pearson correlation to predict orientation specific test performance
500 based on the fitted model parameters of each participant (Fig 4).

501

502 **References:**

- 503 Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit
504 contributions to statistical learning. *Journal of Memory and Language*, *83*, 62–78.
- 505 Brockmole, J. R., & Henderson, J. M. (2006). Using real-world scenes as contextual cues for
506 search. In *Visual Cognition* (Vol. 13, Issue 1, pp. 99–108).
507 <https://doi.org/10.1080/13506280500165188>
- 508 Chukoskie, L., Snider, J., Mozer, M. C., Krauzlis, R. J., & Sejnowski, T. J. (2013). Learning
509 where to look for a hidden target. *Proceedings of the National Academy of Sciences of the*
510 *United States of America*, *110 Suppl 2*, 10438–10445.
- 511 Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current*
512 *Opinion in Neurobiology*, *17*(2), 177–184.
- 513 Findlay, J. M., & Gilchrist, I. D. (2003). *Active Vision: The Psychology of Looking and Seeing*.
514 Oxford University Press.
- 515 Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial
516 structures from visual scenes. *Psychological Science*, *12*(6), 499–504.
- 517 Glimcher, P. W. (2003). The neurobiology of visual-saccadic decision making. In *Annual Review*
518 *of Neuroscience* (Vol. 26, Issue 1, pp. 133–179).
519 <https://doi.org/10.1146/annurev.neuro.26.010302.081134>
- 520 Gottlieb, J. (2012). Attention, learning, and the value of information. *Neuron*, *76*(2), 281–295.
- 521 Greenwald, A. G., & Banaji, M. R. (2017). The implicit revolution: Reconceiving the relation
522 between conscious and unconscious. *The American Psychologist*, *72*(9), 861–871.

- 523 Hannula, D. E., Baym, C. L., Warren, D. E., & Cohen, N. J. (2012). The eyes know: eye
524 movements as a veridical index of memory. *Psychological Science*, *23*(3), 278–287.
- 525 Hannula, D. E., & Ranganath, C. (2009). The eyes have it: hippocampal activity predicts
526 expression of memory in eye movements. *Neuron*, *63*(5), 592–599.
- 527 Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive*
528 *Sciences*, *9*(4), 188–194.
- 529 Henderson, J. M., Hayes, T. R., Rehrig, G., & Ferreira, F. (2018). Meaning Guides Attention
530 during Real-World Scene Description. *Scientific Reports*, *8*(1), 13504.
- 531 Hoppe, D., & Rothkopf, C. A. (2016). Learning rational temporal eye movement strategies.
532 *Proceedings of the National Academy of Sciences of the United States of America*,
533 *113*(29), 8332–8337.
- 534 Hoppe, D., & Rothkopf, C. A. (2019). Multi-step planning of eye movements in visual search.
535 *Scientific Reports*, *9*(1), 144.
- 536 Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10),
537 1295–1306.
- 538 Karuza, E. A., Farmer, T. A., Fine, A. B., Smith, F. X., & Jaeger, T. F. (2014). On-line measures
539 of prediction in a self-paced statistical learning task. *Proceedings of the Annual Meeting of*
540 *the Cognitive Science Society*, *36*(36).
- 541 Kowler, E. (2011). Eye movements: the past 25 years. *Vision Research*, *51*(13), 1457–1483.
- 542 Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning
543 through probabilistic program induction. *Science*, *350*(6266), 1332–1338.
- 544 Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball.
545 *Nature Neuroscience*, *3*(12), 1340–1345.
- 546 Li, C.-L., Pilar Aivar, M., Tong, M. H., & Hayhoe, M. M. (2018). Memory shapes visual search
547 strategies in large-scale environments. In *Scientific Reports* (Vol. 8, Issue 1).
548 <https://doi.org/10.1038/s41598-018-22731-w>

- 549 Mack, S. C., & Eckstein, M. P. (2011). Object co-occurrence serves as a contextual cue to guide
550 and facilitate visual search in a natural viewing environment. *Journal of Vision*, 11(9), 1–16.
- 551 Morvan, C., & Maloney, L. T. (2012). Human visual search does not maximize the post-saccadic
552 probability of identifying targets. *PLoS Computational Biology*, 8(2), e1002342.
- 553 Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search.
554 *Nature*, 434(7031), 387–391.
- 555 Nelson, J. D., & Cottrell, G. W. (2007). A probabilistic model of eye movements in concept
556 formation. *Neurocomputing*, 70(13-15), 2256–2272.
- 557 Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual
558 information and abrupt onsets. *Perception & Psychophysics*, 63(7), 1239–1249.
- 559 Roediger, H. L., 3rd. (1990). Implicit memory. Retention without remembering. *The American*
560 *Psychologist*, 45(9), 1043–1056.
- 561 Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for
562 accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–
563 237.
- 564 Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone
565 sequences by human infants and adults. *Cognition*, 70(1), 27–52.
- 566 Schütz, A. C., Trommershäuser, J., & Gegenfurtner, K. R. (2012). Dynamic integration of
567 information about salience and value for saccadic eye movements. *Proceedings of the*
568 *National Academy of Sciences of the United States of America*, 109(19), 7547–7552.
- 569 Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical
570 learning: Current pitfalls and possible solutions. *Behavior Research Methods*, 49(2), 418–
571 432.
- 572 Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2018). Redefining “Learning” in
573 Statistical Learning: What Does an Online Measure Reveal About the Assimilation of Visual
574 Regularities? *Cognitive Science*, 42 Suppl 3, 692–727.

- 575 Smith, C. N., Hopkins, R. O., & Squire, L. R. (2006). Experience-dependent eye movements,
576 awareness, and hippocampus-dependent memory. *The Journal of Neuroscience: The*
577 *Official Journal of the Society for Neuroscience*, 26(44), 11304–11312.
- 578 Spalek, T. M., & Hammad, S. (2005). The left-to-right bias in inhibition of return is due to the
579 direction of reading. *Psychological Science*, 16(1), 15–18.
- 580 Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical
581 learning. *Journal of Experimental Psychology. General*, 134(4), 552–564.
- 582 Vö, M. L.-H., & Wolfe, J. M. (2013). The interplay of episodic and semantic memory in guiding
583 repeated search in scenes. *Cognition*, 126(2), 198–212.
- 584 Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. In
585 *Nature Human Behaviour* (Vol. 1, Issue 3). <https://doi.org/10.1038/s41562-017-0058>
- 586 Yang, S. C.-H., Lengyel, M., & Wolpert, D. M. (2017). Correction: Active sensing in the
587 categorization of visual patterns. *eLife*, 6. <https://doi.org/10.7554/eLife.25660>
- 588 Yarbus, A. L. (1967). Eye Movements During Perception of Moving Objects. In *Eye Movements*
589 *and Vision* (pp. 159–170). https://doi.org/10.1007/978-1-4899-5379-7_7
- 590