

# Intact Contextual Cueing for Search in Realistic Scenes with Simulated Central or Peripheral Vision Loss

Stefan Pollmann<sup>1-3,\*</sup>, Franziska Geringswald<sup>2,\*</sup>, Ping Wei<sup>1</sup>, and Eleonora Porracin<sup>2</sup>

<sup>1</sup> Beijing Key Laboratory of Learning and Cognition and School of Psychology, Capital Normal University, Beijing, China

<sup>2</sup> Department of Psychology, Otto-von-Guericke-University, Magdeburg, Germany

<sup>3</sup> Center for Behavioral Brain Sciences, Otto-von-Guericke-University, Magdeburg, Germany

**Correspondence:** Stefan Pollmann, Beijing Key Laboratory for Learning and Cognition, School of Psychology, Capital Normal University, Beijing 100048, China. e-mail: [stefan.pollmann@ovgu.de](mailto:stefan.pollmann@ovgu.de)

**Received:** September 20, 2019

**Accepted:** May 29, 2020

**Published:** July 10, 2020

**Keywords:** contextual cueing; real-world scenes; scotoma simulation; gaze-contingent display; explicit memory

**Citation:** Pollmann S, Geringswald F, Wei P, Porracin E. Intact contextual cueing for search in realistic scenes with simulated central or peripheral vision loss. *Trans Vis Sci Tech.* 2020;9(8):15, <https://doi.org/10.1167/tvst.9.8.15>

**Purpose:** Search in repeatedly presented visual search displays can benefit from implicit learning of the display items' spatial configuration. This effect has been named contextual cueing. Previously, contextual cueing was found to be reduced in observers with foveal or peripheral vision loss. Whereas this previous work used symbolic (T among L-shape) search displays with arbitrary configurations, here we investigated search in realistic scenes. Search in meaningful realistic scenes may benefit much more from explicit memory of the target location. We hypothesized that this explicit recall of the target location reduces visuospatial working memory demands on search considerably, thereby enabling efficient search guidance by learnt contextual cues in observers with vision loss.

**Methods:** Two experiments with gaze-contingent scotoma simulation (Experiment 1: central scotoma, Experiment 2: peripheral scotoma) were carried out with normal-sighted observers (total n = 39/40). Observers had to find a cup in pseudorealistic indoor scenes and discriminate the direction of the cup's handle.

**Results:** With both central and peripheral scotoma simulation, contextual cueing was observed in repeatedly presented configurations.

**Conclusions:** The data show that patients suffering from central or peripheral vision loss may benefit more from memory-guided visual search than would be expected from scotoma simulation and patient studies using abstract symbolic search displays.

**Translational Relevance:** In the assessment of visual search in patients with vision loss, semantically meaningless abstract search displays may gain insights into deficient search functions, but more realistic meaningful search scenes are needed to assess whether search deficits can be compensated.

## Introduction

We often encounter situations with a familiar spatial structure. For instance, when we drive towards a street crossing, we expect the road signs or traffic lights in a specific position relative to the road. Often, we learn these spatial relations in an incidental way. When your supermarket suddenly rearranges its shelves, you notice that search becomes awkward, even though you may never have thought about the specific location of a particular shelf before the rearrangement.

In the lab, when a target object is encountered repeatedly in similar spatial contexts, the context-target

configuration can be learned, leading to more efficient search for the target in this context. This contextual cueing effect<sup>1</sup> is usually investigated by presenting visual search displays in which a target needs to be found among a set of distractor items. Typically, one half of the displays presented in the first block is repeated in subsequent blocks, so that the target-distractor configuration can be learned incidentally. The other half of the displays is randomly rearranged in each block, so that no learning of the target-distractor configuration can occur (for methodologic details, see reference 2). Contextual cueing manifests itself in reduced search times for repeated displays relative to new displays.

In previous research, foveal vision loss has been found to eliminate or strongly reduce this contextual cueing, both with central scotoma simulation<sup>3,4</sup> and in patients with age-related macular degeneration (AMD).<sup>5</sup> A closer look revealed that the observed lack of contextual cueing was not due to deficient learning, but rather due to deficient expression of learning, that is, deficient use of learned configurations for efficient search guidance. This could be demonstrated by an immediate search advantage for displays that had been repeatedly presented with gaze-contingent central scotoma simulation when the scotoma simulation was subsequently removed.<sup>4</sup> Previous work has shown that expression of context learning, but not the context learning itself, depends on visuospatial working memory.<sup>6–8</sup> It was hypothesized that the inefficient, top-down controlled search with a central scotoma simulation may leave not enough visuospatial working memory capacity for the expression of context learning.<sup>4</sup>

In these experiments, contextual cueing was investigated in search for a T-shape among L-shaped distractors, a paradigm used in most of the contextual cueing studies since the original paper by Chun and Jiang.<sup>1</sup> In this search context, search facilitation for repeated displays develops gradually with incidental context learning over repetitions.

In contrast, search contexts are typically remembered explicitly in contextual cueing experiments with realistic scenes.<sup>9,10</sup> These experiments, in contrast to search in symbolic displays, have shown an early onset of search facilitation and clear evidence of explicit memory for the repeated target-context configurations. This differed from search in symbolic displays, where participants were only able to recognize a fraction of the repeated displays when explicitly tested<sup>11,12</sup> and the size of search facilitation for repeated displays is typically not different for displays that are explicitly remembered or not remembered.<sup>13</sup> Thus the semantic context inherent in the realistic scenes appears to facilitate explicit learning and retrieval of target-context configurations substantially. A hint that this may also be the case in observers with central vision loss comes from research on object memory: Using a visual object-change detection paradigm, in which participants had to memorize objects in pseudo-realistic scenes comparable to the present study, patients with foveal vision loss due to AMD showed normal sensitivity to subtle object changes.<sup>4</sup>

Not only central vision loss led to deficient contextual cueing in previous studies, but peripheral vision loss as well. In two studies from different labs, search with a simulated peripheral scotoma abolished the search advantage for repeated displays.<sup>4,14</sup> The

results differed in the contextual cueing pattern in a subsequent transfer phase without scotoma simulation. Whereas the pattern observed by Zang et al.<sup>14</sup> indicated intact learning of repeated configurations during search with a peripheral scotoma, the data from Geringswald and Pollmann<sup>4</sup> suggested that no learning of configurations was possible during search with the peripheral scotoma. There were no obvious discrepancies in scotoma size or display parameters that could easily explain the different results. However, in any case these studies jointly showed that contextual cueing was not observed during search with a peripheral scotoma.

In this article, we present two experiments with simulated foveal respectively peripheral vision loss to investigate contextual cueing in repeatedly presented realistic scenes. We hypothesized that the easier encoding or retrieval of the explicitly learned target location in realistic scenes may facilitate contextual cueing in the presence of vision loss—in contrast to search in symbolic displays. In particular, we expected faster manual responses and more efficient visual exploration of the scenes, including a reduced number of fixations and more efficient scan paths, in repeated compared to novel scene target configurations and we expected these contextual cueing effects to be comparable between unconstrained vision and central or peripheral vision impairment. In addition, we expected normal contextual cueing to go along with increased overt visual exploration behavior via eye movements. In particular, we expected an overall increased number of fixations under central vision impairment, in contrast to the reduced number of fixations indicating strongly top-down controlled search that we had previously reported in search in arbitrary T/L-layouts.<sup>3,4</sup>

## Materials and Methods

### Participants

Sixty undergraduate students of the University of Magdeburg participated in this experiment after signing an informed consent form and 20 participants were each randomly assigned to the normal vision (19 females, 1 male, age  $M = 20.25$ ,  $SD = 1.94$ ), scotoma (10 females, 10 males, age  $M = 25.1$ ,  $SD = 3.4$ ) and tunnel vision group (8 males, 12 females, average age 23.6 years,  $SD = 6.04$ ). All participants had self-reported normal or correct-to-normal vision, and they were naive with respect to the purposes of the study. Participants could choose to be compensated with course credits or with a remuneration of 6 Euros/hour. The experiment was carried out in accordance with the

Declaration of Helsinki and was approved by the ethics review board of the University of Magdeburg.

## Apparatus

Stimulus presentation and response recording were performed using a standard PC running Debian Linux and PsychoPy version 1.82.<sup>15</sup> Stimuli were displayed on a 23.6-inch BenQ XL2410T LCD monitor that was 522 mm (1200 pixels) wide and 294 mm (900 pixels) high and had a vertical refresh rate of 120 Hz and has previously been reported to perform very well in experiments that require brief displays.<sup>16</sup> The stimuli were viewed binocularly from a distance of 80 cm and subtended a visual angle of  $36.12^\circ \times 20.82^\circ$  and responses were recorded with a standard computer mouse. The eye position of the left eye was recorded using an Eyelink 1000 Desktop Mount (SR Research Ltd., Mississauga, Ontario, Canada) allowing for remote detection of eye movements, using corneal reflection and pupil tracking, with a temporal resolution of 1,000 Hz. Head movements were minimized by stabilizing participants' heads using a chin and forehead rest.

## Stimuli and Design

We used realistic images of indoor rooms as stimuli in this experiment. All images were created and 3D-rendered with open source interior design software (Sweet Home 3D, version 4.6, eTeks Paris, France, <http://www.sweethome3d.com>) and represented rooms of the following categories: bathroom, bedroom, cinema room, game room, garage, children's room, kitchen, library, living room, concert room, office and study, and each type of room was presented just once in the experiment. The scenes were displayed at a resolution of 1200 x 900 pixels in full color.

Each scene included a yellow cup, which constituted the target in the visual search task. The cup's position was equally allocated to six equal-sized rectangular segments, three (left, middle, right) above respectively below the horizontal midline. The cup was always presented in a meaningful position (e.g. cup on the table, but not floating in the air). The mean width of the cup was  $1.51^\circ$  of visual angle (SD = 0.60) and its height was  $1.18^\circ$  (SD = 0.54) across all scenes. Participants were not informed that some of the scenes and target positions were repeated.

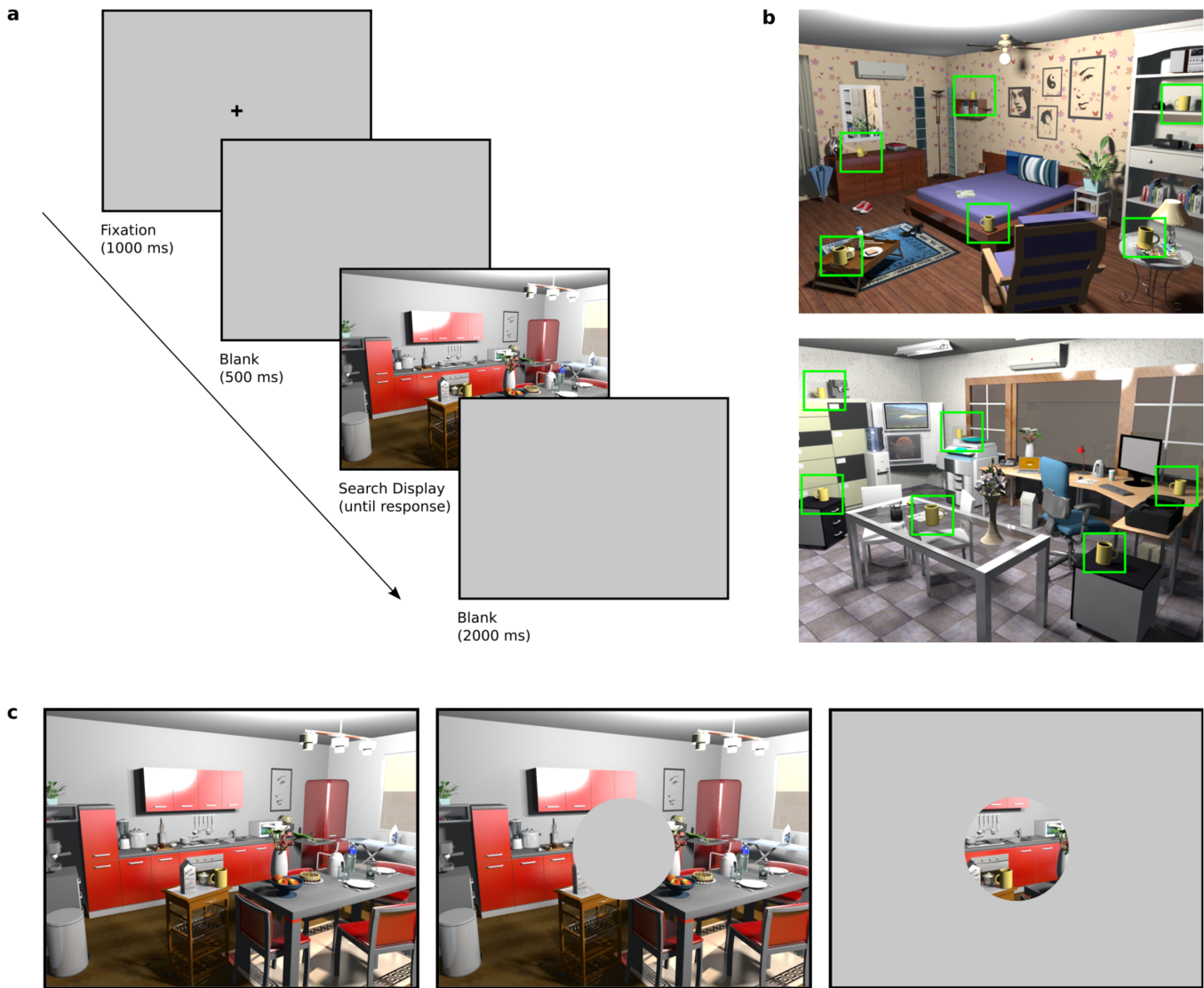
In the experiment, 72 stimuli were presented, divided into six blocks of 12 trials each. Twelve scenes were repeatedly presented across blocks. In six of the scenes, the position of the target was repeated, whereas in the other six scenes the target was randomly

presented in a different display segment in each block. Stimulus sequence was individually randomized for each subject. The repeated scenes were randomly drawn from the pool of all scenes, individually for each participant.

The gaze-contingent central scotoma was created by displaying a solid gray disk with radius of  $5^\circ$  of visual angle, covering the foveal and parafoveal region of the retina (moving mask technique).<sup>17</sup> Inversely, the gaze-contingent tunnel window (moving window paradigm)<sup>18</sup> with a radius of  $5^\circ$  visual angle allowed exclusively for foveal and parafoveal vision. The position of the gaze-contingent window was continuously updated with the coordinates (x,y) provided by the eye-tracker device throughout each search trial (gaze-contingent protocol)<sup>19,20</sup>, and the gaze samples were filtered by the heuristic one-sample filter<sup>21</sup> implemented in the Eyelink software, removing single-sample noise artifacts. The Eyelink 1000 average end-to-end delay was 2.8 ms, and the worst-case latency until the update of the gaze-contingent stimulus manipulation on the monitor was two frames (16.7 ms). No other additional filter algorithms—for example, for fixation or saccade identification—were implemented. The estimated worst-case delay between actual gaze position and stimulus update was thus about 20 ms. When the gaze coordinates were unavailable because of eye blinks or signal losses, the gaze-contingent aperture remained on the last measured valid gaze position until a new gaze sample became available. Please note that this minimal processing of the gaze samples may allow for transient motion of the scotoma at the beginning and end of eye blinks and does not prevent the unwanted triggering of slow eye movements by the scotoma as a moving target stimulus itself (see Aguilar and Castet<sup>20</sup> for a more advanced approach).

## Procedure

Participants searched for a yellow cup within a realistic scene and indicated the position of its handle on the left or the right side of the cup with left or right mouse button presses, respectively. Participants were told to respond as quickly and accurately as possible to promote a passive search strategy as suggested by Lleras and von Mühlhelen.<sup>22</sup> In the artificial scotoma conditions, a gaze-contingent central or peripheral scotoma was presented throughout the search task. The experiment consisted of the contextual cueing task comprising six blocks of 12 trials each and the subsequent recognition task comprising six trials that were performed by all participants. Participants of the control and scotoma condition performed a subsequent scotoma validation task in addition.

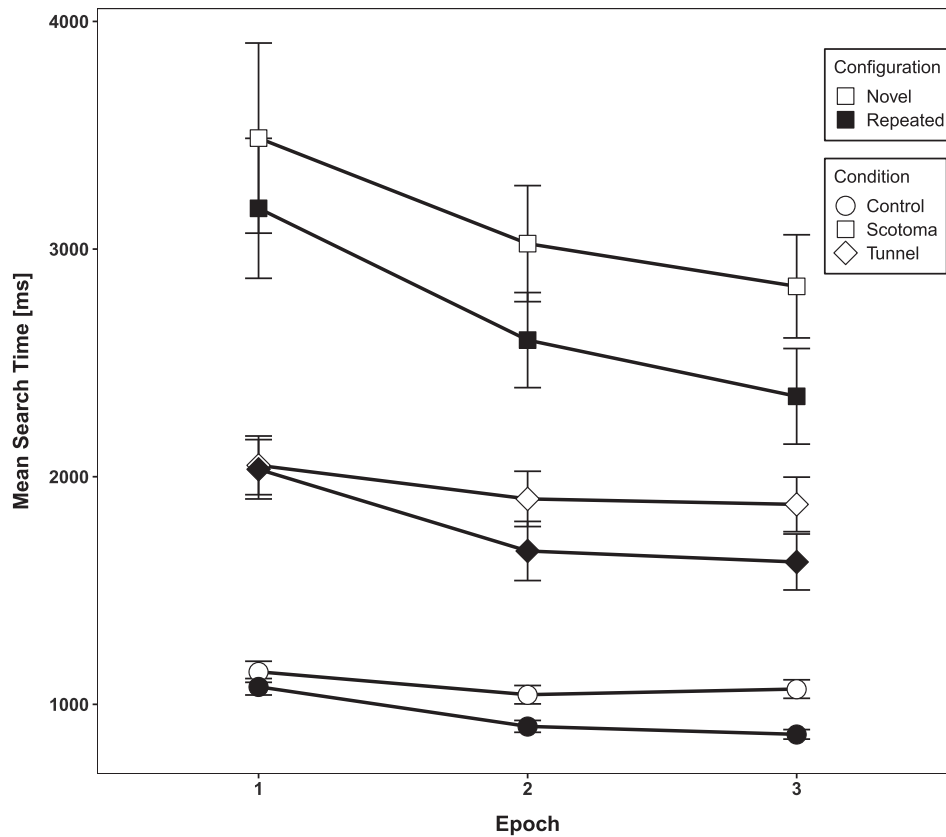


**Figure 1.** Visualization of the stimuli and simulated scotomata. (a) Schema of an experimental trial of the search task. Each trial consisted of a fixation cross (1000 ms), followed by a blank screen (500 ms), the search scene (presented until response), and a blank screen (2000 ms). (b) Exemplary search scenes and target positions. Each scene contained one yellow target cup. In repeated displays, a scene was always paired with a specific target position during the experiment. In novel displays, the target appeared once at each of the predefined positions across experimental blocks. The *green squares* are for illustrative purposes and were not presented during the experiment. (c) Experimental vision conditions. The control group searched for the cup without vision impairment (left), the scotoma group searched with a gaze-contingent central scotoma extending across foveal and parafoveal vision with a radius of 5° visual angle (middle) and the tunnel group perceived the displays through a moving window with a radius of 5° (right).

Each trial of the contextual cueing task started with a fixation cross that was presented at the center of the screen in black on a gray background for 1000 ms (Fig. 1). Participants were instructed to fixate the cross. After the fixation cross, the search display was presented and remained visible until the response. After the response a feedback sound (a high pitch tone for the correct response and a low pitch tone for the wrong response) was provided. Before each experimental block, a nine-point gaze calibration was performed,

followed by a nine-point gaze validation. The calibration and validation procedure was repeated in case the error was bigger than 1° on average or bigger than 1.5° for the worst point. The gaze-contingent display was presented neither during the calibration nor the validation procedure.

After a short break, all participants completed the recognition task in which they viewed the six real-world rooms for which the target location was repeated during the first part of the experiment. These scenes



**Figure 2.** Averaged search times for the viewing conditions control (*circles*), scotoma (*squares*), and tunnel (*diamond*) separated for repeated (*filled symbols*) and novel displays (*open symbols*). Error bars depict the standard error of the mean.

were rendered without the search target (i.e., the yellow cup). At the beginning of each trial, a fixation cross was presented at the center of a gray background. Participants were asked to move the mouse cursor to the location they believed to have seen the cup most often and confirm this location with a mouse click. No feedback was provided.

The scotoma validation task was based on a search-discrimination task that uses Landolt C stimuli and requires high-acuity foveal vision in normal viewing.<sup>23</sup> The task consisted of 4 blocks with 32 stimuli each, in the first and third block participants searched without scotoma simulation, in the second and fourth block participants searched with simulated central scotoma (solid gray disk with a radius of 5° visual angle as in the main experiment). At the beginning of each trial, a fixation cross was presented at the center of the screen for 1000 ms, followed by a Landolt C (i.e. a circle with a gap at the left, bottom, right or top) among seven same-sized rings used as distractors. These stimuli were placed on an invisible circle with a radius of 10° visual angle. The stimulus remained on the screen for 5000 ms or until the response. After the response a feedback sound (a high pitch tone for correct

responses, a low pitch tone for incorrect responses) was provided. The size of the Landolt C was chosen so that the opening (extending 0.067°) should not be discernible with peripheral vision in the presence of the central scotoma (see 23 for details).

## Data Analysis

All statistical tests were carried out using R (version 3.5.1;<sup>24</sup>). The six experimental blocks were aggregated to three epochs to increase statistical power. Analyses of variance (ANOVAs) were performed using type III sums of squares. For all statistical tests, the alpha level was set to .05. Contextual cueing was analyzed with three-way mixed-design ANOVAs with the between-subjects factor viewing condition (control, central scotoma, peripheral scotoma) and the within-subjects factors epoch<sup>1-3</sup> and configuration (repeated, novel). Analyses were based on log-transformed RT to account for the large baseline differences between the scotoma group and the full vision group (Fig. 2). Overall differences in exploration behavior were analyzed with one-way ANOVAs with the between-subjects factor vision group using only novel configurations of the last epoch

of the contextual cueing task, when observers had become accustomed to the scotoma simulations. All post hoc *t*-tests were two-tailed and corrected according to Holm.<sup>25</sup>

Gaze parameters were identified with the Eyelink Dataviewer Software (SR Research Ltd., Mississauga, Ontario, Canada), using velocity and acceleration thresholds of 35°/s and 9500°/s<sup>2</sup>, respectively, for saccade detection. To analyze contextual cueing effects, we calculated the number of fixations required to find the target, the efficiency of the scan path toward the target as the ratio between the total distance covered by the eye during search, and the shortest path possible and the onset of the monotonic gaze path, characterizing the onset of a more direct path toward the target location after an initial inefficient search phase, as dependent measures. In typical contextual cueing experiments, observers show less fixations, more efficient scan paths and an earlier onset of the monotonic gaze approach in repeated configurations, accompanying speeded response times.<sup>3,26–29</sup> To further characterize differences exploration behavior between the vision groups, the average number and duration of fixations and saccade amplitudes were calculated as dependent measures.

For the analysis of the recognition task, the (x,y) coordinates of each response was first scored as correct or incorrect. Responses were defined as correct when the (x,y) coordinates fell within the image segment in which the target cup was originally presented during search in the respective repeated display. Performance values were then averaged for each participant and analyzed with a one-way ANOVA with the between-subjects factor vision group. In addition, recognition performance of each group was compared against chance level with one-sample *t* tests. Since all scenes were divided into six equal-sized rectangular segments that could contain the target in the contextual cueing task, chance performance in the recognition task was thus 1 out of six (16%).

## Data Exclusion

After the initial period of data collection, one participant of the scotoma group (excessive errors, 23.6%) and one of the tunnel vision group (unstable gaze tracking) were replaced with new participants who received the exact same stimulus sequences respectively. To ensure that the gaze-contingent scotoma simulations were presented reliably during search, we removed all search trials in which the amount of signal loss exceeded a maximum threshold of 20%. This affected, on average, 0.07% (SD = 0.31%), 0.42% (SD = 0.79%), and 1.38% (SD = 0.31%) of trials in

the control, scotoma, and tunnel vision group. Search times and gaze parameters were only analyzed for correct responses. We removed trials with search times exceeding 30 seconds or the mean plus two standard deviations as individual cut-off, affecting on average 4.27% (SD = 1.40%), 8.67% (SD = 2.80%), and 4.56% (SD = 2.05%) of trials in the control, scotoma, and tunnel vision groups. Gaze data of one participant in the tunnel vision group were not available for analysis.

## Results

### Accuracy

Overall accuracy during the search task was very high in all vision groups. A repeated-measures ANOVA with vision (control, scotoma, tunnel) as between-subjects factor, and configuration (repeated, novel) and epoch<sup>1–3</sup> as within-subject factors revealed a significant effect of vision ( $F[2, 57] = 18.86, P < 0.001, \eta^2_P = 0.398, \eta^2_G = 0.177$ ). This was due to the somewhat lower accuracy in the scotoma group (93.80%) than in the control (99.44%,  $P < 0.001$ ) and tunnel vision group (98.23,  $P < .001$ ). None of the other effects approached significance (all  $F$ s  $< 2.13, P$ s  $> 0.15, \eta^2_P < 0.041, \eta^2_G < 0.009$ ).

### Contextual Cueing

The repeated-measures ANOVA with vision (control, scotoma, tunnel) as between-subjects factor, and configuration (repeated, novel) and epoch<sup>1–3</sup> as within-subject factors on log-transformed RT to account for the large baseline differences between vision groups (Fig. 2) revealed a significant main effect of vision ( $F[2, 57] = 4.50, P < 0.05, \eta^2_P = 0.136, \eta^2_G = 0.134$ ), reflecting significantly longer search times with central scotoma simulation (2913 ms) than in the full vision group (1016 ms,  $P < 0.05$ ). There was a nonsignificant trend towards longer search times in the tunnel vision group compared to controls (1860 ms,  $P = 0.08$ ) and search time was not significantly different in the artificial vision groups ( $P = 0.43$ ). Moreover, significant main effects for epoch ( $F[2, 114] = 41.96, P < 0.001, \eta^2_P = 0.424, \eta^2_G = 0.007$ ) and configuration ( $F[1, 57] = 32.23, P < 0.001, \eta^2_P = 0.361, \eta^2_G = 0.004$ ) were observed, indicating faster search over time as well as faster search in repeated (1813 ms) compared with novel displays (2048 ms). The epoch by configuration interaction was significant as well ( $F[2, 114] = 2.26, P < .01, \eta^2_P = 0.109, \eta^2_G = 0.001$ ), indicating that contextual cueing developed over time (131, 264, and

**Table 1.** Mean Difference Between Novel and Repeated Configurations and Normalized Contextual Cueing Effects ((RT<sub>novel</sub>-RT<sub>repeated</sub>)/RT<sub>novel</sub>)

Measure	Condition	Epoch 1		Epoch 2		Epoch 3	
		Absolute Mean (SD)	Normalized Mean (SD)	Absolute Mean (SD)	Normalized Mean (SD)	Absolute Mean (SD)	Normalized Mean (SD)
Search time (ms)	Control	66 (148)	0.04 (0.13)	139 (160)	0.11 (0.14)	199 (164)	0.16 (0.13)
	Scotoma	308 (1294)	-0.07 (0.30)	424 (922)	0.06 (0.31)	483 (1136)	0.09 (0.37)
	Tunnel	17 (443)	-0.03 (0.23)	229 (516)	0.09 (0.27)	253 (337)	0.12 (0.18)
Number of fixations (counts)	Control	0.175 (0.57)	0.01 (0.17)	0.59 (0.69)	0.12 (0.20)	0.81 (0.80)	0.15 (0.22)
	Scotoma	0.81 (3.68)	-0.09 (0.35)	1.49 (3.05)	0.07 (0.42)	1.46 (3.01)	0.10 (0.35)
	Tunnel	-0.05 (1.92)	-0.06 (0.30)	1.00 (2.15)	0.07 (0.32)	1.19 (1.24)	0.13 (0.18)
Scan pattern ratio (quotient)	Control	0.07 (0.50)	-0.13 (0.46)	0.41 (0.44)	0.13 (0.40)	0.73 (1.04)	0.14 (0.55)
	Scotoma	0.08 (5.53)	-0.37 (0.67)	2.96 (10.23)	-0.12 (0.86)	1.54 (5.37)	-0.15 (0.76)
	Tunnel	0.07 (0.67)	-0.13 (0.51)	0.50 (0.78)	0.10 (0.51)	0.38 (0.58)	0.11 (0.32)
Onset of the monotonic scan path (counts)	Control	-0.04 (0.06)	-0.10 (0.24)	0.35 (0.69)	0.02 (0.35)	0.64 (0.77)	0.16 (0.25)
	Scotoma	0.81 (3.86)	-0.12 (0.39)	1.64 (3.08)	0.08 (0.49)	1.33 (2.97)	0.09 (0.38)
	Tunnel	0.16 (1.56)	-0.07 (0.31)	0.73 (2.48)	-0.04 (0.58)	0.68 (1.48)	0.02 (0.30)

**Table 2.** Statistical Results of the Between-Group Analyses of Contextual Cueing for Number of Fixations, Scan Pattern Ratio and Monotonicity of the Scan Path

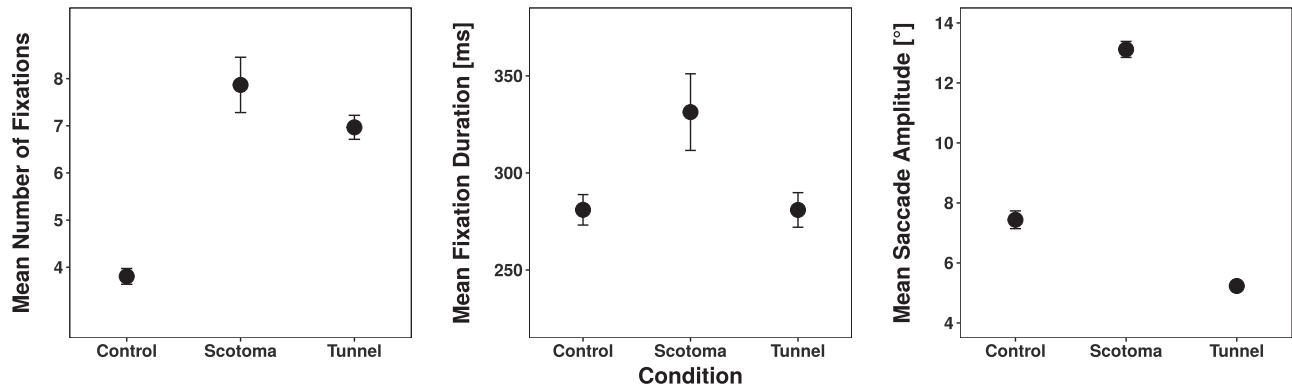
Effect	Measure													
			Number of fixations				Scan Pattern Ratio				First fixation of Monotonic Path			
	DFn	DFd	F	P	$\eta^2_P$	$\eta^2_G$	F	P	$\eta^2_P$	$\eta^2_G$	F	P	$\eta^2_P$	$\eta^2_G$
Vision	2	56	46.94	<0.001	.6263	.4669	52.71	<0.001	.6531	.4253	53.86	<0.001	.6580	.4823
Epoch	2	112	16.38	<0.001	.2263	.0494	0.80	0.45	.0142	.0029	14.95	<0.001	.2108	.0489
Configuration	1	56	19.93	<0.001	.2625	.0431	3.17	0.09	.0536	.0124	14.00	<0.001	.2000	.0320
Vision × Epoch	4	112	1.40	0.24	.0475	.0088	0.89	0.47	.0309	.0064	1.58	0.18	.0535	.0108
Vision × Configuration	2	56	1.40	0.25	.0477	.0063	0.87	0.43	.0301	.0069	2.34	0.11	.0772	.0109
Epoch × Configuration	2	112	2.89	0.06	.0491	.0089	1.79	0.17	.0309	.0057	1.50	0.23	.0261	.0051
Vision × Epoch × Configuration	4	112	0.18	0.95	.0065	.0011	0.85	0.50	.0295	.0055	0.11	0.98	.0038	.0007

312 ms in first, second, and third epochs, respectively; see Table 1). Crucially, the other interactions containing the factor vision were not significant (all  $F_s < 0.90$ ,  $P_s > 0.46$ ,  $\eta^2_P < 0.031$ ,  $\eta^2_G < 0.001$ ), indicating that the contextual cueing effect and its development was comparable between vision groups.

The results of the repeated measures ANOVAs on gaze parameters are summarized in Table 2 and the mean differences between repeated and novel configurations can be found in Table 1. All gaze parameters indicated the complicated search with the gaze-contingent simulations as indicated by the significant main effects of vision. The number of fixations ( $F[2, 56] = 46.94$ ,  $P < 0.001$ ,  $\eta^2_P = 0.626$ ,  $\eta^2_G = 0.467$ ) was significantly increased in the scotoma (7.87 fixations) and tunnel vision groups (6.80 fixations) compared with the control group (3.60 fixations,  $P_s < 0.001$ ) and the scotoma simulation also led to significantly more fixations than the tunnel simulation ( $P < 0.05$ ). The scan pattern ratio ( $F[2, 56] = 52.71$ ,  $P < 0.001$ ,  $\eta^2_P = 0.653$ ,  $\eta^2_G = 0.425$ ) was significantly impaired by the simulated scotoma (7.55) compared with

controls (1.28,  $P < 0.001$ ) and tunnel vision (1.73,  $P < 0.001$ ) while the efficiency of scan paths was comparable between controls and tunnel vision ( $P = 0.51$ ). The onset of the monotonic path ( $F[2, 56] = 53.86$ ,  $P < 0.001$ ,  $\eta^2_P = 0.658$ ,  $\eta^2_G = 0.482$ ) was significantly delayed by the scotoma (7.15 fixations) and tunnel vision groups (5.14 fixations) compared with controls (2.64 fixations,  $P_s < 0.001$ ), and the scotoma simulation also delayed the onset more than the tunnel simulation ( $P < 0.001$ ).

The effects of contextual cueing were most prominent in the number of fixations as indicated by a significant main effect of configuration ( $F[1, 56] = 19.93$ ,  $P < 0.001$ ,  $\eta^2_P = 0.263$ ,  $\eta^2_G = 0.043$ ), reflecting a reduced number of fixations in repeated (5.66) compared to novel configurations (6.49), a significant main effect of epoch ( $F[2, 112] = 16.38$ ,  $P < 0.001$ ,  $\eta^2_P = 0.226$ ,  $\eta^2_G = 0.049$ ) reflecting overall improvement over time, and a non-significant trend in the configuration by epoch interaction ( $F[2, 112] = 2.89$ ,  $P = 0.06$ ,  $\eta^2_P = 0.049$ ,  $\eta^2_G = 0.009$ ). Similarly, contextual cueing was indicated by the significant main



**Figure 3.** Averaged number of fixations (left), fixation duration (middle), and saccade amplitude (right) of novel configurations in the last epoch as a function of vision condition. *Error bars* depict the standard error of the mean.

effect of configuration ( $F[1, 56] = 14.00, P < 0.001, \eta^2_P = 0.200, \eta^2_G = 0.032$ ) in the onset of the monotonic path, reflected by an earlier onset of the monotonic path in repeated (4.62 fixations) compared with novel configurations (5.33 fixations) and a non-significant trend of the main effect configuration in the scan pattern ratios ( $F[1, 56] = 3.17, P = 0.08, \eta^2_P = 0.054, \eta^2_G = 0.012$ ). None of the interactions containing vision group reached significance (all  $F_s < 2.34, P_s > 0.10, \eta^2_P < 0.078, \eta^2_G < 0.010$ ), indicating that contextual cueing, when observable in the gaze patterns, did not differ between vision groups.

## Visual Exploration

To further investigate the influence of the gaze contingent simulations on visual behavior, we performed one-way ANOVAs with the between-subjects factor vision on the number of fixation, fixation duration and saccade amplitude. Only novel configurations of the last epoch were included to test exploration behavior when observers were maximally accustomed to the gaze-contingent display manipulations (Fig. 3). All effects were significant (all  $F_s > 4.72, P_s < 0.05, \eta^2_P > 0.144, \eta^2_G > 0.144$ ). Search with the scotoma led to an increased number of fixations (7.86 fixations,  $P < 0.001$ ) with a significantly longer duration (331 ms,  $P < 0.05$ ) and significantly increased saccade amplitudes (13.12° visual angle,  $P < 0.001$ ) than in controls (number of fixations: 3.81, fixation duration: 281 ms, saccade amplitude: 7.44° visual angle). The tunnel simulation also led to a significant increase in fixation numbers (6.97 fixations,  $P < 0.001$ ), but did not affect fixation duration (281 ms,  $P = 1$ ) and

saccade amplitudes were reduced (5.23° visual angle,  $P < 0.001$ ) compared with controls.

## Recognition Task

A one-way ANOVA with the between-subjects factor vision on the recognition performance revealed a significant effect ( $F[1, 57] = 3.48, P < 0.05, \eta^2_P = 0.109, \eta^2_G = 0.109$ ), qualified by significantly higher recognition accuracy in the control group (71.67%,  $SD = 21.01\%$ ) than in the scotoma group (53.33%,  $SD = 24.54\%$ ,  $P < 0.05$ ). Recognition accuracy amounted to 62.50% ( $SD = 20.14\%$ ) in the tunnel group and was not significantly different from the other two groups (all  $P_s > 0.38$ ). To test whether participants were able to recall the position of the target within each vision group, we performed additional one-sample  $t$ -tests against chance. In each trial, the probability to select the right position segment out of the six display segments was 1/6 and the chance level value was thus set to 16%. All  $t$ -tests were significant (all  $t[19] > 6.68, P_s < 0.001$ ), indicating that all participants, independent of the vision condition, recalled the target location in the repeated displays.

## Scotoma Validation Task

Without scotoma, on average, the direction of the gap was identified correctly in 60.65 of 64 trials (95%), whereas in the condition with the scotoma, only 19.95 correct responses of 64 (31%) were recorded. The performance dropped significantly ( $t[19] = 20.78; P < 0.001$ ), reflecting an impairment in the stimulus discrimination, demonstrating that vision was compromised by the scotoma simulation.



## Discussion

In previous experiments, foveal vision loss, both natural and simulated, led to a severe reduction or complete loss of contextual cueing in symbolic (T among L) search.<sup>4,5</sup> Likewise, gaze-contingent simulated peripheral vision loss, forcing tunnel vision, also abolished contextual cueing.<sup>4,14</sup> In contrast, in the present experiments, investigating search in realistic scenes, contextual cueing was unimpaired during central, as well as peripheral scotoma simulation.

Central scotoma simulation interfered substantially with visual search, indicated by the three- to fourfold longer search times. Longer search times were accompanied by complicated visual exploration of the scenes, reflected by a roughly doubled number of fixations required to find the target. Moreover, the efficiency of the scotoma simulation was independently confirmed with a behavioral validation task. Nevertheless, there was a strong advantage for repeated displays searched with central scotoma simulation that was evident in manual response times and more efficient visual exploration of repeated configurations, in stark contrast to the absence of contextual cueing in T-among-L search with foveal vision loss.<sup>3-5</sup>

Brockmole et al.<sup>9</sup> showed that the global scene layout is of particular importance for contextual cueing in realistic scenes. Contextual cueing survived local changes in repeated displays but was severely reduced when the global scene layout was changed and only the local target context remained unchanged. In contrast, repetition of only the local configuration led to comparable search facilitation than repetition of the full display in T-among-L search,<sup>30</sup> although gross changes of the global context<sup>30</sup> or search with peripheral scotoma simulation<sup>4,14</sup> also eliminated contextual cueing in these symbolic displays. The stronger impact of the local configuration on contextual cueing may at least in part be due to the high target-distractor similarity and the low between-distractor similarity in T-among-L search that makes distractor grouping difficult.<sup>31</sup> In comparison, natural scenes are typically much more structured, so that the target object can more easily be related to the scene layout. Realistic scenes can be easily remembered, so that hundreds of scenes can be recognized after a single presentation,<sup>32,33</sup> whereas only a fraction of the arbitrary symbolic configurations often used in contextual cueing tasks can be remembered, even after multiple repetitions.<sup>12</sup> In addition to the better structured scene layout, the better memory for scenes apparently rests on the categorization of the scene (kitchen, sleeping room, etc.) with an easy semantic description of the position of individual items in

the respective room (e.g., vase on the table). Brockmole and Henderson<sup>26</sup> demonstrated the facilitation of search by semantic scene context in that mirror images of repeated scenes still facilitated search. This semantic representation along with the explicit memory of the target location in the repeated scene enables the observer a more efficient search for the target than in repeated arbitrary displays.

Previous findings showed that contextual cueing in symbolic displays suffered under foveal vision loss due to impaired memory-driven search guidance. This could be demonstrated in that contextual cueing was immediately present when scotoma simulation was removed in a test phase following search with the scotoma.<sup>4</sup> This indicated that expression of learning (i.e., search guidance by repeated displays), rather than learning of the repeated contexts, suffered from foveal vision loss. Search guidance by repeated contexts has been shown to depend on visuospatial working memory resources.<sup>6-8</sup> Consequently, we have hypothesized that the need for top-down controlled eye movement planning in search with central scotoma simulation may compete for visuospatial working memory resources, thereby limiting resources available for memory-driven search guidance.<sup>4</sup> A significantly reduced number of fixations in search with a central scotoma supported this hypothesis, indicating a more top-down controlled slow search with a potentially increased attentional focus.<sup>4</sup> This limitation appears to be less problematic for search in realistic displays, because of the explicit memory of the target location and the generally more efficient search for the target in realistic scenes. In contrast to arbitrary search layouts, search with the central scotoma lead to an increased amount of fixations and the difference in fixation duration was less pronounced than in our previous studies, proposing a more natural search mode in realistic scenes.

The scotoma size was chosen to encompass foveal and parafoveal vision and was very similar to our previous simulations (5° radius vs. 4.5° radius in Geringswald and Pollmann<sup>4</sup>). In contrast to our previous study, we chose a hard edge of the scotoma here. Thus the present scotoma simulation covered even a bit more space than the previous one that created difficulties with contextual cueing in the T-among-L search paradigm. Nevertheless, contextual cueing remained intact, and we do therefore not believe that the soft edge played a significant role in disturbing contextual cueing in our previous work. A more subtle simulation of the scotoma may make visual exploration more cumbersome and interfere with contextual cueing in natural scenes, however. We have previously shown that both a solid and a

semitransparent scotoma mask led to highly similar exploration of natural scenes with respect to fixation number and duration as well as saccade amplitude.<sup>34</sup> This suggests that contextual cueing may remain intact, even when the borders of the central scotoma are less visible as in patients with naturally occurring scotomata. Nevertheless, follow-up experiments with scotoma patients may be valuable to assess the generality of our present finding of intact contextual cueing in the presence of scotomatous vision. Our past work, however, showed a good agreement between scotoma simulation and patient studies, for instance in contextual cueing studies with symbolic displays.<sup>4-5</sup>

Simulation of a gaze-contingent scotoma is not a trivial task, especially in complex natural stimuli. Patients are often unaware of their scotomata, potentially due to perceptual filling-in.<sup>35</sup> In more artificial stimuli, a similar negative effect of a central scotoma can simply be simulated by using the same color for the scotoma and the background. In this way, the scotoma can only be perceived when it covers a patch of a distinct color. Increased visibility of the scotoma on natural stimuli, on the other hand, bears a greater risk of unwanted side effects of the simulation. For example, undesired motion of the simulation due to eye blinks or slow eye movements caused by the tracking of the scotoma itself can interfere with experimental factors. Due to the minimal processing of the gaze coordinates used for the scotoma simulation, these effects may not be accounted for sufficiently in the current study (for a detailed demonstration of these issues and their prevention, see Aguilar and Castet<sup>20</sup>). The impact of these factors may be limited in our setup; however, because we used an Eyelink 1000 desktop mount with a refresh rate of 1000 Hz that uses more robust algorithms to reduce the impact of blink-induced changes in pupil shape on gaze location, a high screen refresh rate of 120 Hz, a relatively large scotoma size of 10° in diameter and relatively large target objects extending about 1°. We do acknowledge, however, that the pitfalls detailed by the article by Aguilar and Castet<sup>20</sup> are an important issue concerning any study that uses gaze-contingent window simulation that should be taken into consideration by future studies.

## Conclusions

We have shown that participants with simulated central or peripheral vision loss search for objects in realistic scenes more efficiently when these scenes are repeatedly presented. This contextual cueing effect was not observed in previous studies using semanti-

cally meaningless symbolic search displays. Contextual cueing in realistic scenes went along with explicit recognition of the target location. Thus, patients with vision loss may be still able to take advantage of repeatedly encountered scenes for search guidance, at least when they are able to access an explicit memory of the scene.

More generally, these data show that it is worthwhile to test if findings from the psychological laboratory hold up in more realistic environments, particularly if the goal is to develop tools to assist patients in making the most out of the information that is available to them, e.g. patients suffering from vision loss that can process only a fraction of the visual field. Realistic scenes often allow the use of alternative strategies based on richer information than the one offered in tightly controlled lab situations.

## Acknowledgments

We thank Isabel Dombrowe for her support in data analysis and Sascha Purmann for his support in experimental programming.

Supported by the Deutsche Forschungsgemeinschaft (PO548/14-2).

Disclosure: **S. Pollmann**, None; **F. Geringswald**, None; **P. Wei**, None; **E. Porracin**, None

\* SP and FG contributed equally to this article.

## References

1. Chun MM, Jiang Y. Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cogn Psychol*. 1998;36:28–71.
2. Jiang YV, Sisk CA. Contextual cueing. In: Pollmann S, ed. *Spatial Learning and Attention Guidance*. Neuromethods, vol. 151. New York, NY: Humana; 2019.
3. Geringswald F, Baumgartner F, Pollmann S. Simulated loss of foveal vision eliminates visual search advantage in repeated displays. *Front Hum Neurosci*. 2012;6:134.
4. Geringswald F, Pollmann S. Central and peripheral vision loss differentially affects contextual cueing in visual search. *J Exp Psychol Learn Mem Cogn*. 2015;41:1485–1496.
5. Geringswald F, Herbig A, Hoffmann MB, Pollmann S. Contextual cueing impairment in patients with age-related macular degeneration. *J Vis*. 2013;13:28–28.

6. Annac E, Manginelli AA, Pollmann S, Shi Z, Müller HJ, Geyer T. Memory under pressure: Secondary-task effects on contextual cueing of visual search. *J Vis.* 2013;13:1–15.
7. Manginelli AA, Geringswald F, Pollmann S. Visual search facilitation in repeated displays depends on visuospatial working memory. *Exp Psychol.* 2012;59:47–54.
8. Manginelli AA, Langer N, Klose D, Pollmann S. Contextual cueing under working memory load: selective interference of visuospatial load with expression of learning. *Atten Percept Psychophys.* 2013;75:1103–1117.
9. Brockmole JR, Castelhana MS, Henderson JM. Contextual cueing in naturalistic scenes: Global and local contexts. *J Exp Psychol Learn Mem Cogn.* 2006;32:699–706.
10. Brockmole JR, Henderson JM. Using real-world scenes as contextual cues for search. *Vis Cogn.* 2006;13:99–108.
11. Geyer T, Baumgartner FJ, Mueller HJ, Pollmann S. Medial temporal lobe-dependent repetition suppression and enhancement due to implicit vs. explicit processing of individual repeated search displays. *Front Hum Neurosci.* 2012;6:272.
12. Geyer T, Shi Z, Müller HJ. Contextual cueing in multiconjunction visual search is dependent on color-and configuration-based intertrial contingencies. *J Exp Psychol Hum Percept Perform.* 2010;36:515–532.
13. Colagiuri B, Livesey EJ. Contextual cuing as a form of nonconscious learning: Theoretical and empirical analysis in large and very large samples. *Psychon Bull Rev.* 2016:1–14.
14. Zang X, Jia L, Müller HJ, Shi Z. Invariant spatial context is learned but not retrieved in gaze-contingent tunnel-view search. *J Exp Psychol Learn Mem Cogn.* 2015;41:807–819.
15. Peirce JW, Gray JR, Simpson S, MacAskill MR, Höchenberger R, Sogo H, Kastman E, Lindeløv J. PsychoPy2: experiments in behavior made easy. *Behavior Res Methods.* 2019;51:195–203.
16. Lagroix HE, Yanko MR, Spalek TM. LCDs are better: Psychophysical and photometric estimates of the temporal characteristics of CRT and LCD monitors. *Atten Percept Psychophys.* 2012;74:1033–1041.
17. Rayner K, Bertera JH. Reading without a fovea. *Science.* 1979;206:468–469.
18. McConkie GW, Rayner K. The span of the effective stimulus during a fixation in reading. *Percept Psychophys.* 1975;17:578–586.
19. Duchowski AT, Cournia N, Murphy H. Gaze-contingent displays: a review. *CyberPsychol Behav.* 2004;7:621–634.
20. Aguilar C, Castet E. Gaze-contingent simulation of retinopathy: some potential pitfalls and remedies. *Vis Res.* 2011;51:997–1012.
21. Stampe DM. Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behav Res Meth Instrum Comput.* 1993;25:137–142.
22. Lleras A, Von Muhlenen A. Spatial context and top-down strategies in visual search. *Spatial Vis.* 2004;17:465–482.
23. Geringswald F, Baumgartner FJ, Pollmann S. A behavioral task for the validation of a gaze-contingent simulated scotoma. *Behav Res Methods.* 2013;45:1313–1321.
24. R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. 2018. Available at: URL <https://www.R-project.org/>. Accessed May 30, 2016.
25. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat.* 1979;6:65–70.
26. Brockmole JR, Henderson JM. Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *Q J Exp Psychol.* 2006;59:1177–1187.
27. Manginelli AA, Pollmann S. Misleading contextual cues: How do they affect visual search? *Psychol Res.* 2009;73:212–221.
28. Peterson MS, Kramer AF. Attentional guidance of the eyes by contextual information and abrupt onsets. *Percept Psychophys.* 2001;63:1239–1249.
29. Tseng Y-C, Li C-SR. Oculomotor correlates of context-guided learning in visual search. *Percept Psychophys.* 2004;66:1363–1378.
30. Brady TF, Chun MM. Spatial constraints on learning in visual search: modeling contextual cuing. *J Exp Psychol: Hum Percept Perform.* 2007;33:798–815.
31. Duncan J, Humphreys GW. Visual search and stimulus similarity. *Psychol Rev.* 1989;96:433–458.
32. Standing L. Learning 10000 pictures. *Quart J Exp Psychol.* 1973;25:207–222.
33. Brady TF, Konkle T, Alvarez GA, Oliva A. Visual long-term memory has a massive storage capacity for object details. *Proc Natl Acad Sci.* 2008;105:14325–14329.
34. Geringswald F, Porracin E, Pollmann S. Impairment of visual memory for objects in natural scenes by simulated central scotomata. *Journal of Vision.* 2016;16:6, doi:10.1167/16.2.6.
35. Zur D, Ullman S. Filling-in of retinal scotomas. *Vis Res.* 2003;43:971–982, [https://doi.org/10.1016/S0042-6989\(03\)00038-5](https://doi.org/10.1016/S0042-6989(03)00038-5).