Visual reinforcement shapes eye movements in visual search

Céline Paeye*

Abteilung Allgemeine Psychologie,
Justus-Liebig-Universität Gießen, Gießen, Germany
Laboratoire Psychologie de la Perception,
Université Paris Descartes, Paris, France ✉

Alexander C. Schütz*

Abteilung Allgemeine Psychologie,
Justus-Liebig-Universität Gießen, Gießen, Germany
Allgemeine und Biologische Psychologie,
Philipps-Universität Marburg, Marburg, Germany ⌂ ✉

Karl R. Gegenfurtner

Abteilung Allgemeine Psychologie,
Justus-Liebig-Universität Gießen, Gießen, Germany ⌂ ✉

We use eye movements to gain information about our visual environment; this information can indirectly be used to affect the environment. Whereas eye movements are affected by explicit rewards such as points or money, it is not clear whether the information gained by finding a hidden target has a similar reward value. Here we tested whether finding a visual target can reinforce eye movements in visual search performed in a noise background, which conforms to natural scene statistics and contains a large number of possible target locations. First we tested whether presenting the target more often in one specific quadrant would modify eye movement search behavior. Surprisingly, participants did not learn to search for the target more often in high probability areas. Presumably, participants could not learn the reward structure of the environment. In two subsequent experiments we used a gaze-contingent display to gain full control over the reinforcement schedule. The target was presented more often after saccades into a specific quadrant or a specific direction. The proportions of saccades meeting the reinforcement criteria increased considerably, and participants matched their search behavior to the relative reinforcement rates of targets. Reinforcement learning seems to serve as the mechanism to optimize search behavior with respect to the statistics of the task.

Introduction

Visual search is a frequently used behavior that includes many psychological aspects, such as visual perception, attention, memory, learning, and decision making (Nakayama & Martini, 2011). Evidence drawn from the literature about these processes shows that low-level properties of the stimuli (Itti & Koch, 2000; Parkhurst, Law, & Niebur, 2002; Wolfe, 2007) and higher level factors interact to guide target selection during visual search (Schütz, Braun, & Gegenfurtner, 2011; Tatler, Hayhoe, Land, & Ballard, 2011). Among the latter top-down influences, task demands, knowledge about the visual properties of the stimuli and statistical regularities of the environment have been shown to contribute to search performance (see Eckstein, 2011, for a review). The visual system also takes into account its own properties when planning fixations, for example the inhomogeneity of the retina (Najemnik & Geisler, 2005, 2008). It is important to understand how this knowledge about the environment and the structure of the visual system is acquired. Here we explore the hypothesis that the reward system, which is generally implicated in the learning of behaviors (Montague, Hyman, & Cohen, 2004) offers a suitable way for implementing the learning process for optimally deploying gaze during visual search tasks.

Several paradigms have been used to investigate the influence of reward on eye movement behavior at different stages of target selection. Both stimulus salience and reward interact to determine the final eye position when participants have to choose between two (Chen, Mihalas, Niebur, & Stuphorn, 2013) or more (Ackermann & Landy, 2013; Eckstein, Schoonveld, Zhang, Mack, & Akbas, 2015; Navalpakkam, Koch, Rangel, & Perona, 2010; Towal, Mormann, & Koch, 2013) items. In these studies saccadic choice is well

described by models dynamically combining these two factors.

However, there are several restrictions in these previous studies. First, saccades were constrained by few possible item locations in the displays. Moreover, these displays were presented for short durations which allowed participants to make only one or two saccades. As a consequence, the tasks were more similar to a selection task than to a free exploration of the environment. Second, the nonvisual nature of saccades' consequences also differentiates visual search as studied in the aforementioned studies from visual search in a natural environment. Whereas an explicit reinforcement, such as points, monetary gains (Tatler et al., 2011), or alimentary reward in monkeys (Hikosaka, Nakamura, & Nakahara, 2006), allows researchers to experimentally manipulate the consequences of eye movements, such reinforcement does not happen in everyday life after making a saccade. Instead, we gain visual information that can indirectly be used to act and get reward from our environment.

Indeed, theoretical models of human search behavior assume that fixation locations are chosen to maximize the information gain across successive eye movements (Najemnik & Geisler, 2005, 2008) and to minimize the uncertainty about the target location (Renninger, Verghese, & Coughlan, 2007) or regions of task-relevant information (Peterson & Eckstein, 2014). In these models the visual information gain can be conceptualized as a rewarding consequence, controlling fixation locations. A few paradigms explicitly evaluated the reinforcing value of *visual* consequences of saccades. However they were simplistic in that observers had to choose between only two visual stimuli presented at predetermined locations (Berlyne, 1972; Collins, 2012) or only investigated the very basic aspects of saccades such as their latency, speed, and amplitudes in monkeys (Dorris, Pare, & Munoz, 2000) as well as in humans (Collins, 2012; Madelain, Paeye, & Wallman, 2011; Montagnini & Chelazzi, 2005; Paeye & Madelain, 2011; Schütz, Kerzel, & Souto, 2014; Xu-Wilson, Zee, & Shadmehr, 2009).

Overall, the picture emerges that eye movements for visual search are optimal or close to optimal under some conditions (Ackermann & Landy, 2013; Clarke, Green, Chantler, & Hunt, 2016; Droll, Abbey, & Eckstein, 2009; Eckstein et al., 2015; Najemnik & Geisler, 2005, 2008), but not others (Morvan & Maloney, 2012; Verghese, 2010). The reason for the different outcomes is unclear at present. We wanted to explore eye movement strategies in a relatively simple paradigm, where correct target location is biased in one particular region of the search display. Efficient search should then also be biased toward this region, because the visibility of the search target is increased in the close vicinity to the fovea. Some previous studies have

addressed this issue. The results indicate that the visual system takes into account prior knowledge about the statistical target distribution (Jiang, Swallow, Rosenbaum, & Herzig, 2013; Jones & Kaschak, 2012; Peterson & Kramer, 2001). However, these studies used simple displays with a small number of potential target locations. In the extreme case, Chukoskie, Snider, Mozer, Krauzlis, and Sejnowski (2013) used a single invisible target, forcing the observers to use their prior experience of reward exclusively. They used tones in order to signal the observers that they "found" the target—and no visual information. Once observers figured out the correct location, they very quickly saccade to that point, a behavior that departs from a visually guided search task.

To circumvent these issues, we used a continuous search display consisting of a 1/f random noise background and Gabor targets with a well-defined contrast. We will first show that the positional bias did not have any effect on search strategies, most likely because the reward structure of the task did not induce reinforcement learning. In a second experiment, we systematically controlled reinforcement rate in a gaze-contingent paradigm, and found a tight coupling between eye movement positional biases and reinforcement rate.

Materials and methods

Participants

Six (five females, one male; aged 21 to 28 years), seven (five females, two males; aged 21 to 30 years) and six (three females, three males; aged 23 to 27 years) participants took part in experiments 1 to 3, respectively. They were students of the University of Giessen and naïve as to the purpose of the study. They had normal or corrected-to-normal vision. They came to the laboratory for several daily one-hour sessions (see Table 1 in Appendix) during which several blocks of 50 trials (separated by 5-min breaks) were recorded. Experiments were in accordance with the principles of the Declaration of Helsinki and approved by the local ethics committee LEK FB06 at the University of Giessen (proposal number 2009-0008). Participants gave informed written consent prior to the experiment. They received eight Euros per hour.

Experiment 1: Frequency biases

In this experiment participants were asked to look for a Gabor patch in a circular 1/f noise background in which we manipulated the likelihood of the target in

different parts of the display. First we measured the foveal contrast sensitivity of each participant with a procedure adapted from the two-interval forced choice paradigm implemented by Najemnik and Geisler (2005, 2008). We measured the detection accuracy only for the center of the display and for the only background contrast used in all our experiments. This preliminary session consisted in 40 trials of a quest procedure (from the Psychophysics Toolbox, Watson & Pelli, 1983). In these trials participants were asked to fixate continually the center of the display (within a $1.5° \times 1.5°$ of visual angle window). A first background noise texture was displayed for 250 ms, followed by a gray screen. After 500 ms a second noise texture appeared for 250 ms. Only one of the two background noise textures contained a Gabor patch (whose contrast depended on the quest procedure) at the central fixation location. Participants judged which interval contained the target. We fitted psychometric functions with a cumulative Gaussian. We could then estimate and adjust the target contrast for each participant individually to achieve a given perceptual performance (d' between 1 and 4). The target root-mean-square contrast varied between 0.05 and 0.12. Targets were presented from the beginning of the trials at one of 104 homogeneously distributed target locations. To induce learning they were located in one of 30 possible locations of one quadrant (the "rich quadrant") four times more often than in all the other 74 possible locations of the search display. Since the rich quadrant also included all locations along the vertical and horizontal midline the chance proportion was not exactly 0.25, but 30/104, i.e., 0.288. Observers were not informed about the manipulation of probability. The number of 50-trials blocks performed by the participants (see Table 1 in Appendix) depended on the time needed to reach a stable performance during three consecutive blocks. Each trial ended when a target was localized.

Experiments 2 and 3: Reinforcement procedures

In the two subsequent experiments we designed a gaze-contingent paradigm in order to provide more direct visual consequences—the target appearance—after saccades meeting specific position (experiment 2) or direction (experiment 3) criteria and according to specific probabilities. Observers were not informed about these reinforcement procedures.

We first measured baseline search behavior during 15 trials in which no target was presented. Each of these trials was cancelled automatically after the execution of ten saccades. Then the learning blocks began. No target was visible at the beginning of each trial. Visual reinforcement was provided on a saccade-to-saccade basis, according to two concurrent schedules. In one of

these schedules, saccades landing in one quadrant of the background (experiment 2) or saccades with a direction located in a specific 60° angle range (experiment 3) were reinforced with a high probability–the quadrants or angle ranges of this high-probability reinforcement alternative were randomly assigned to each participant. At the same time according to the other reinforcement schedule, saccades in the three other quadrants (experiment 2) or with other directions (experiment 3) were reinforced with a low probability. Table 1 (Appendix) presents these scheduled reinforcement probabilities and the rates of reinforcement actually obtained by each participant (which depended first on our ability to predict saccadic vectors and second on the detection of the target by the participants). Depending on individual performance, the probabilities of reinforcement could be modified during the experiment, but the programmed ratio remained constant. This was the case for Participant FR (experiment 2): Given her quick learning pace (her proportion of saccades landing in the highly reinforced quadrant was above 80% during the two last blocks of her first daily session), we decreased the reinforcement rate in this quadrant from 0.8 to 0.4—and from 0.2 to 0.1 in the three other quadrants—for her second session. Similarly, in experiment 3, the proportion of saccades we programmed to reinforce according to their direction decreased from 100% to 50% from the second session of participant FL. Finally, following two daily sessions, three participants in experiment 2 and one in experiment 3 performed a second experimental condition: Unknown to them, the reinforcement criteria were changed and another quadrant (experiment 2) or another 60° angle range (experiment 3) defined the high-probability reinforcement schedules.

Eye movement recording

Eye movements were recorded with a video-based eye tracker (EyeLink 1000; SR Research, Kanata, Ontario, Canada) and were sampled at 1000 Hz. For offline analysis, we used the Eyelink parser to identify saccades' onset and offset, using 30° of visual angle per second (°/s) velocity and $8000°/s^2$ acceleration thresholds.

Presenting the target right at the end of a saccade requires knowledge of the landing position while the saccade is still in midflight. We used the stereotyped relationship between saccadic amplitude, peak velocity, and duration (Bahill, Clark, & Stark, 1975) to make this prediction (Figure 1A).

First, we predicted the saccadic peak velocity. Each millisecond, a buffer of 50 eye velocity samples was updated, and the corresponding velocity trace was fitted with a Gaussian (gray curve in the example
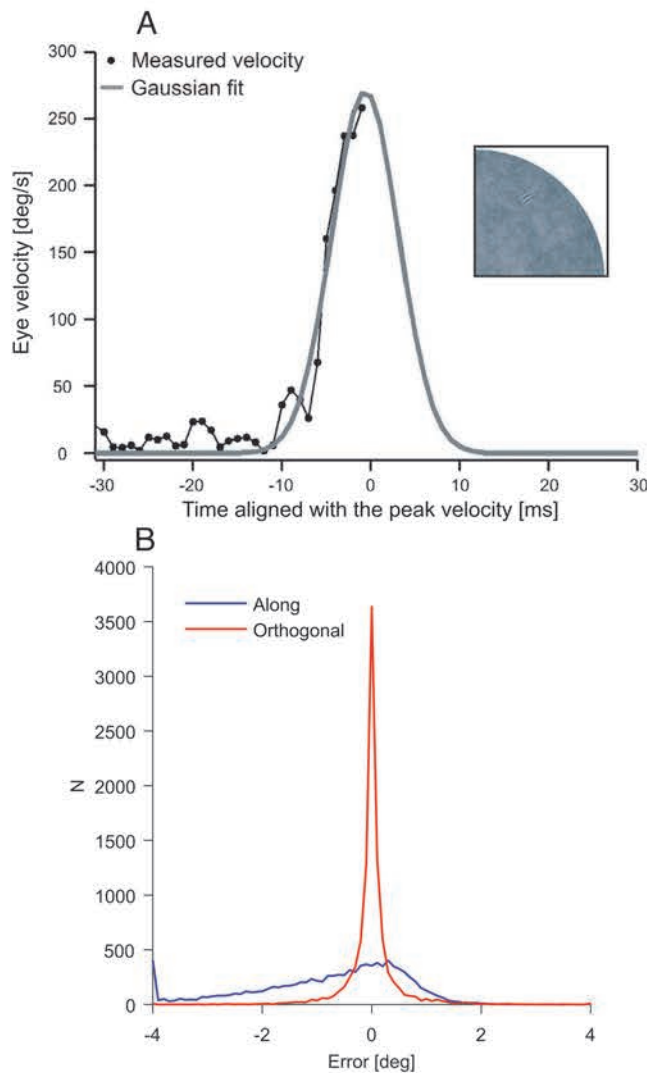
Figure 1. Saccade endpoint prediction: (A) Example of the last 30 ms out of the 50 stored in the buffer used to predict a saccadic peak velocity. Each millisecond, this buffer was updated and the velocity trace was fitted with a Gaussian (gray curve). The predicted peak velocity was computed from this fit as soon as its maximum exceeded 150°/s while the standard deviation remained below 20 (as illustrated here). The predicted landing position was used to present the target in the noisy background (inset) at saccade offset after a saccade met the reinforcement criteria. (B) Histogram of the prediction error along the movement trajectory (blue) and orthogonal to the movement direction (red).

presented in Figure 1A; the black symbols correspond to the last 30 samples obtained at the time a landing position was predicted). Once the maximum of the Gaussian was larger than 150°/s and the standard deviation smaller than 20, the peak velocity was determined. In addition the mean of the Gaussian had to be located within −10 to 2 ms relative to the current sample. With these criteria the online detection rate of saccades was between 29% for saccades with ampli-

tudes of 2° and 57% for saccades with amplitudes of 9°. Although the detection rate was higher for saccades with larger amplitudes, saccade amplitudes did not increase over the course of experiment 2, first trials: $M = 3.4°$, $SD = 0.7°$; last trials: $M = 2.7°$, $SD = 0.6°$; $t(6) = 1.82$; $p = 0.119$; or experiment 3, first trials: $M = 3.4°$, $SD = 0.9°$; last trials: $M = 3.2°$, $SD = 0.6°$; $t(5) = 0.83$; $p = 0.446$.

If the above mentioned criteria were met, the saccade amplitude was predicted with the following formula, using the maximum (max), the mean, and the standard deviation (*SD*) of the Gaussian fit:

$$-1.32 + \text{max} \times 0.16 + SD \times 0.01 + \text{mean} \times 0.04$$
$$+ \frac{SD}{\text{max}} \times 1.06 \tag{1}$$

These values have been optimized in advance using offline measurements of saccade trajectories. The saccade direction was predicted as the average direction between the onset and the peak of the saccade. If the estimated vector of the current saccade met the reinforcement criteria, the target appeared at the predicted landing position.

We were able to present the target with a median delay of 5 ms in experiment 2 and three ms in experiment 3 after saccade offset and with a median distance of respectively 1.09° and 1.36° from the actual eye landing positions (Figure 1B). While saccade direction was estimated accurately, saccade amplitude was slightly underestimated (median = −0.44°). For the vast majority of trials the prediction error was small, so that the target was presented in the foveal region of the retina with the highest acuity and sensitivity.

Stimuli and experimental procedure

Stimuli were generated using the Psychophysics Toolbox (Kleiner, Brainard, & Pelli, 2007) for Matlab and displayed on a video monitor (Samsung, driven at 120 Hz). The participants were seated in a dark room, and their heads were stabilized by a chin and forehead rest at 80 cm from the screen. We used stimuli analogous to those of Najemnik and Geisler (2005, 2008). The search target was a sine-wave grating of 6 cycles/°, tilted 45° to the right. In the gaze-contingent paradigms, the target contrast was fixed with a root-mean-square of 0.30 (see inset in Figure 1A). It was displayed on a 0.07 root-mean-square contrast background, a circular region 15° in diameter filled with 1/f noise at a mean luminance of 46 cd/m², a condition which resembles the spatial statistics of natural scenes.

At the beginning of each experiment, participants were given the instruction "to look for the Gabor patch in the circular gray background," and they could freely explore an example of the display containing an

embedded target. Prior to each block (consisting of 50 trials), a 9-point-grid calibration procedure was applied. Then a fixation cross (0.5° × 0.5°, line width 2 pixels, either black or white) appeared at the center of the screen, and the participant pressed one of the buttons of a joystick to begin the trial, which extinguished the fixation cross. As soon as participants detected the target, they pressed a second joystick button, and a large white "plus" appeared (line width 3 pixels), which covered the whole screen and whose center was moving along with the measured eye position. The participants indicated the target location by looking at the target (which brought the cross center on the target) and pressing the first joystick button.

Behavior analysis

We had to determine for each trial, which saccade was actually being rewarded, i.e., which saccade led to target detection by the participant. Indeed, in the frequency biases experiment, participants could have looked at locations close to the target without noticing it. In the gaze-contingent paradigm, they might also have missed the target when it was first presented. We only considered saccades that ended after the target presentation and before the first button press, which indicated that the participant found the target. If more than one saccade satisfied these criteria (83% of trials in experiment 1, 52% of trials in experiment 2, and 62% of trials in experiment 3), we calculated for each potential saccade the distance to the target and how much it reduced this distance. Saccades were only considered if the distance to the target location was below 2° after this or after the following saccade and if the saccade reduced the distance by at least 2°. The first saccade that satisfied these criteria was selected as the rewarded saccade. The results were qualitatively similar if only the last matching saccade was selected. In more than half of the trials (56% of trials in experiment 1, 66% of trials in experiment 2, and 64% of trials in experiment 3), only one saccade matched these criteria.

To describe the modifications in visual search behavior, we assessed the changes in the proportion of saccades landing in the rich quadrant (experiments 1 and 2) or the proportion of saccades whose direction was located in the frequently reinforced angle range (experiment 3). These proportions were computed for each participant and for each trial over a moving window of 40 trials as well as over specific blocks of trials: baseline trials, the last 100 trials of each experimental condition. Paired *t* tests were used to compare baseline and final trials across all participants and chi-square (χ^2) tests were used to compare baseline and final trials for individual participants. A significance criterion of 0.05 was used for all statistical tests.

To analyze the relationship between eye movements and targets' detection, we used the generalized matching equation proposed by Baum (1974) to formalize the relationship between behavior and reinforcement:

$$\log\left(\frac{S_F}{S_O}\right) = s \times \log\left(\frac{R_F}{R_O}\right) + b. \qquad (2)$$

Here, S corresponds to the number of saccades made in the rich quadrant or in the frequently reinforced direction ("F") and in the other areas or directions ("O"). R is the number of reinforcers received (i.e., targets detected) after saccades meeting the high- and low-probability reinforcement criteria. In this linear equation, a slope of $s = 1$ indicates that relative response rates will perfectly match relative reinforcement rates produced by this behavior. The intercept b gives a measure of the preference for one response, independently of the reinforcement rates. We fitted the data obtained over the last 100 saccades of each experimental condition.

A matching analysis was also used to disentangle the effects of our two reinforcement procedures. We reanalyzed the data obtained in the quadrant experiment in terms of saccadic directions. Reciprocally, we reanalyzed the data of the direction experiment in terms of locations. For example, in the latter experiment, a saccade starting from the screen center and made in a 10 o'clock direction would be coded as landing in the upper left background's pie-shaped sector, whereas a saccade with the same vector but starting from the bottom of the background would be coded as landing in a sector situated below the horizontal median line. We assumed that the rich pie-shaped sector of the search display corresponded to the frequently reinforced angle bin.

Results

We wanted to explore the effect of positional biases of the search target on eye movement strategies in unconstrained visual search. If observers adapt their eye movements to these biases, they should have more fixations in the biased quadrant.

Experiment 1: Frequency biases

In this conventional, nongaze contingent, visual search task, one of the quadrants of the display contained the target four times more often than the others. Participants made 6 (*SD* = 1.82) saccades per trial on average, which shows that the target could not be easily detected in the periphery.
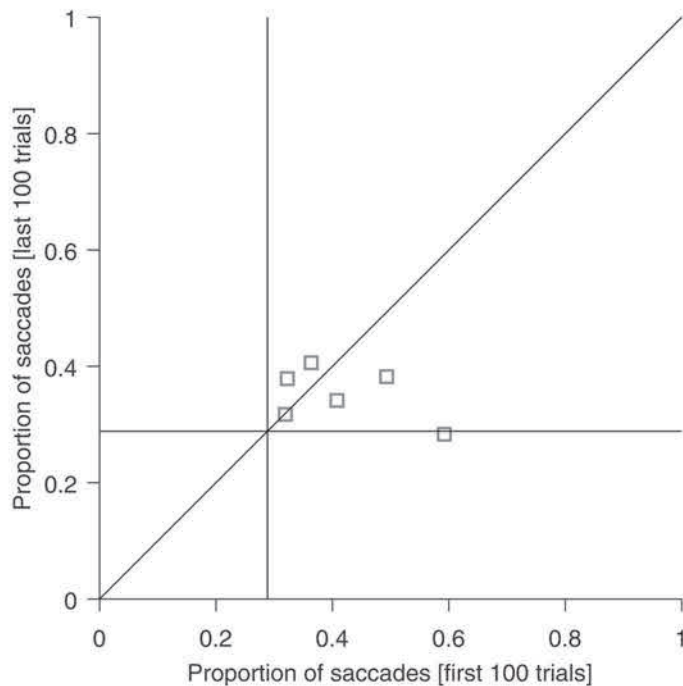
Figure 2. Results of the frequency biases experiment. Proportions of saccades made in the high target probability quadrant for all participants. These proportions were computed over the first 100 trials and plotted against these proportions computed over the last 100 trials. Horizontal and vertical lines: chance proportions.

Figure 2 shows that observers overall exhibit a slight bias towards the rich quadrant. However, this bias did not change between the first ($M = 42\%$; $SD = 11\%$) and last 100 trials ($M = 35\%$; $SD = 5\%$), indicating that the bias was constant over the course of the experiment, $t(5) = 1.14$, $p = 0.305$. There was also no change if we considered only the first saccades in each trial; first trials: $M = 53\%$, $SD = 24\%$; last trials: $M = 54\%$, $SD = 21\%$; $t(5) = 0.25$; $p = 0.810$. For each two of our subjects, the bias increased significantly, decreased significantly, or remained constant (individual χ^2 tests). Thus, observers seem to be unable to acquire useful information about the inhomogeneity of the reward landscape over the course of the experiment. The general bias towards the rich quadrant was most likely due to the visual signal that made the search target visible when the eyes were in the vicinity.

To understand why observers failed to exploit the statistics of the environment, we looked at the relationship between the reinforcement history and the selection of saccade target locations. Indeed, in learning experiments, participants often closely match reinforcement rates in their behavior (Herrnstein, 1961). In this first experiment, the reward was pre determined because the target was present in the display independently of observer's eye movements. Participants' failure to modify their search behavior might be due to the fact that the statistical bias we introduced in our display was too small.

Despite the particular target bias of 4:1 we had chosen, the reward ratios—i.e., the number of rewarded saccades in the rich quadrant versus those in the lean quadrants—were rather uniform and low; and this was reflected in the saccade ratios—i.e., the number of all saccades in the rich quadrant versus those in the lean quadrants (Figure 3A). In other words, in this conventional search task, the large number of possible target locations made it difficult for the observers to learn the reward structure of the environment. Since the target locations were deterministic, we can estimate the theoretical reward ratios for different biases towards the rich quadrant, assuming that target detection is homogeneous across the search display and that targets are predominantly detected by saccades landing in the same quadrant. This analysis of the theoretically possible reward ratios (Figure 3B) reveals that huge biases and a large number of trials would be necessary to create high reward ratios. In previous studies, this problem did not arise because of the very limited number of distinct potential target locations, which allow high reward ratios even with smaller target biases and less repetitions.

Experiment 2: Reinforcement of saccade landing positions

Our first experiment aimed at testing whether the visual system would modify its search behavior if one part of the display contained significantly more often the target. Surprisingly, we found that participants did not search more often in the biased area. We argue that the ratio of target detection in the rich quadrant and the lean quadrants was too low to modify search behavior. We address this issue here by introducing a gaze-contingent search display. The search target was displayed at the landing position of the saccade, when saccades were predicted to land in the reinforced quadrant. This procedure gives us complete control over the reinforcement schedule, despite the complexity of the stimulus and the task.

According to the visual reinforcement procedure, the target was presented after saccades landing in a specific quadrant (the "rich quadrant") four times more often than after saccades made in the other quadrants.

Figure 4A shows data from a representative participant. Over the course of the experiment, the percentage of saccades landing in the rich quadrant increased from the baseline at 20% to 60%. Following the modification of the reinforcement criteria and the specification of a new rich quadrant, the proportions of saccades landing in this new rich quadrant increased from 11% to 86%. At the same time they decreased in
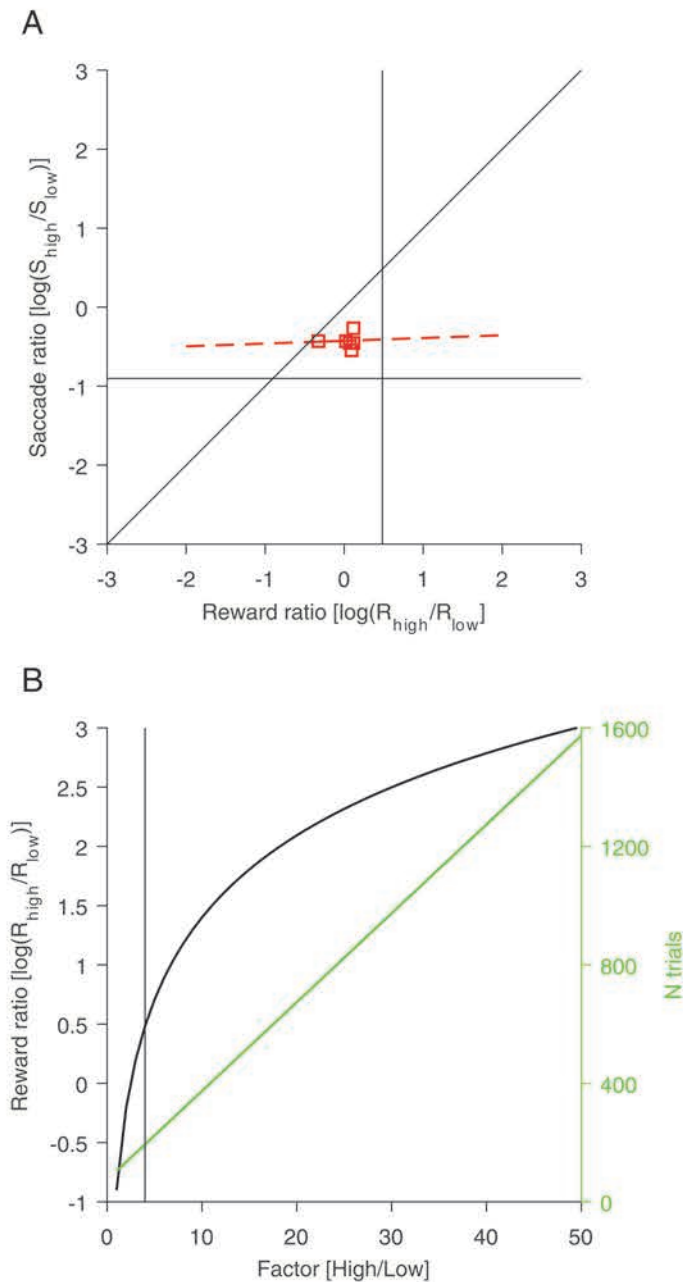
A

B

Figure 3. Matching analysis. (A) Log ratio of saccades made in the high target probability quadrant and low target probability quadrants plotted against the log ratio of number of saccades that led to target detection in these areas. Each data point represents data of one observer obtained over the last 100 trials. Red line: best fit line according to the generalized matching equation (Baum, 1974). The line is solid in the range of measured reward ratios and dashed outside this range. Black diagonal line: perfect matching. Black horizontal line: theoretical saccade ratio if saccades were equally distributed. Black vertical line: theoretical reward ratio given all possible target locations. (B) Analysis of theoretically possible reward ratios (black) and resulting trial numbers (green) as a function of how often targets were repeated in the rich quadrant. The vertical line represents the empirically tested factor of 4. The

the previously rich quadrant down to 3%. Thus, reinforcement of saccades landing in a certain quadrant led to an overall increase of saccades landing in that quadrant.

Figure 4B summarizes the results obtained for all seven participants of this experiment. All of them except one modified their search behavior and looked preferentially for the target in the quadrant where saccades were most frequently followed by a target presentation. The proportion of saccades landing in this quadrant nearly tripled, increasing from 21% (SD 4%) to 60% (SD 23%) on average between the baseline trials and the end of the first condition. The increase was significant for the whole group of participants, $t(6) = 5.01$, $p = 0.002$, as well as for six out of the seven individual participants, $\chi^2(1)$ tests, all $p < 0.001$. A similar increase from 13% (SD 11%) to 76% (SD 37%) was present if we considered only the first saccade in each trial, $t(6) = 5.48$, $p = 0.002$. One participant did not show any signs of learning (but see matching analysis below). For the three participants who performed a second condition with another rich quadrant, the proportion increased from 4% (SD = 6%) to 63% (SD = 33%) on average, individual $\chi^2(1)$ tests, all $p < 0.001$. This indicates that gaze-contingent reinforcement is very effective in changing visual search behavior.

It is notable in Figure 4B that there were substantial interindividual differences in the proportion of saccades to the rich quadrant, ranging from 13% (the participant who did not learn) to 89%. One possible reason is that due to the individual learning history the actual reinforcement rates were different for the individual participants (see Table 1 in Appendix). Another potential cause for differences is that our algorithm did not perfectly detect saccades, thereby slightly affecting the actual reinforcement rates.

It is instructive to use the matching equation as a model to study the relation between saccade landing positions and detected targets. The model fit (red line in Figure 5) accounted remarkably well for the data ($R^2 = 0.94$). The slope of the regression line was close to unity (1.21; 95% CI [0.97; 1.45]), indicating that matching was observed across participants. The analysis also points to a bias towards the quadrants associated with a low probability of reinforcement, since the intercept

←

calculations are based on 104 target locations overall and 30 target locations in the rich quadrant. Each target in the low-probability quadrants is shown once, and the factor specifies how often targets in the rich quadrant are shown. This prediction requires the assumption that target detection is homogeneous across the search display, i.e., identical for rich and lean quadrants and that targets are detected only by saccades landing in the same quadrant.
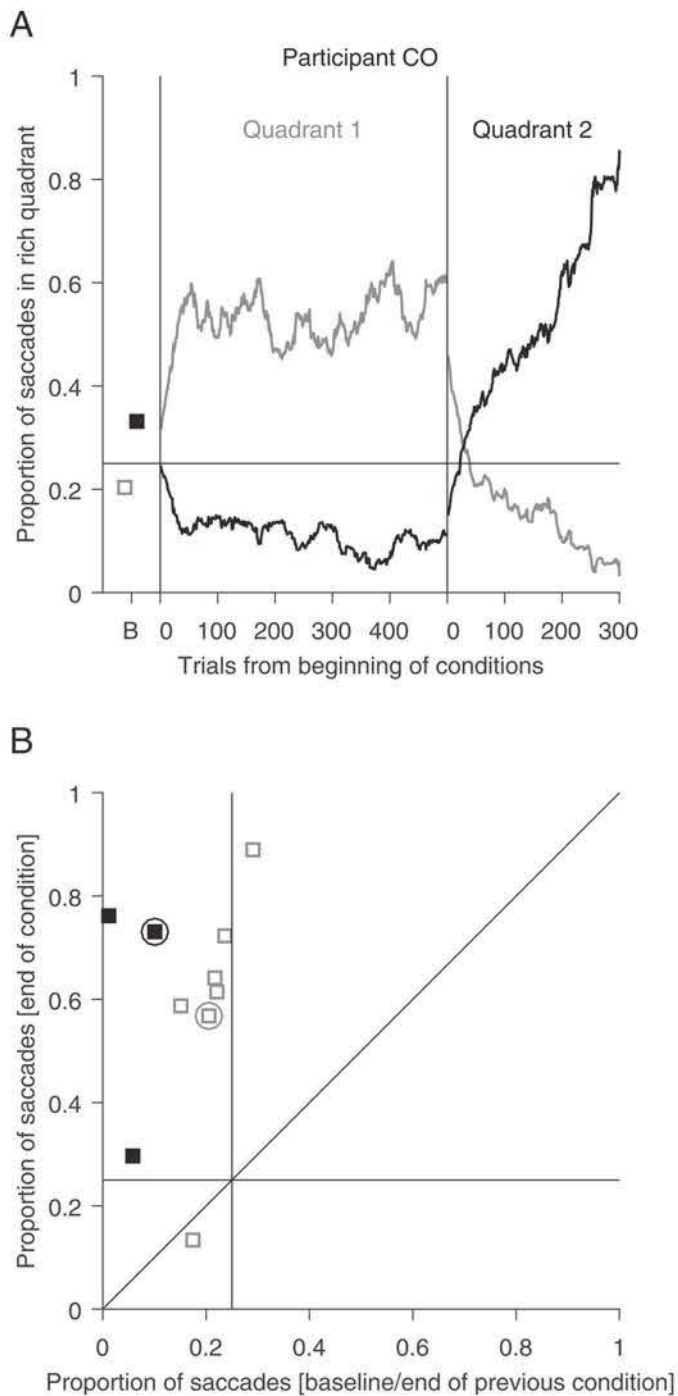
Figure 4. Location experiment. (A) Individual results: evolution of the proportion of saccades meeting the criteria of the high-probability reinforcement schedules for participant CO. For each trial, this proportion was computed over a moving window encompassing 40 trials. Horizontal line: chance proportion. In both conditions we programmed to reinforce 40% of the saccades landing in a specific rich quadrant versus 10% of the saccades landing in the three other quadrants. The open square and the gray curve correspond to the quadrant initially reinforced, the filled square and the black curve to the quadrant reinforced after the change in the reinforcement schedule. (B) Global results: Proportions of saccades meeting the criteria of

of the line was significantly below zero (−1.10; 95% CI [−1.45; −0.75]).

Curiously, the search strategy of one participant (MA) could not be modified by our reinforcement procedure, even after more than 500 trials and despite an actual reinforcement rate in the rich quadrant three times higher than in the other quadrants (see Table 1 in Appendix). We observed that this participant persevered in the same strategy consisting in scanning the search background in successive horizontal lines in the reading direction, starting from the upper-left corner. The application of such a rigid search-strategy led to an extremely low reward ratio, which impeded the learning of reinforcement contingencies. This is illustrated in Figure 5: The data point from this observer, at the bottom of the figure, is located on the best fit according to the matching equation.

These results show that direct visual consequences do not only lead to changes in search behavior. It is the individual reinforcement history that determines the rate of change. In contrast to the first experiment, our gaze-contingent paradigm allows achieving sufficiently high reinforcement rates to effectively modify the observers' eye movements.

Experiment 3: Reinforcement of saccade directions

In this experiment we probed the extent of control we can exert on visual search behavior by using a visual reinforcing consequence. We tested whether finding the search target could reinforce saccade directions rather than landing positions. Then we performed two matching analyses to determine whether the search behavior effectively followed a reinforcement schedule of saccade directions or if it was still controlled by the reinforcement of specific landing positions. In this experiment the rich schedule consisted in presenting the target after 50% or 100% of the saccades moving at an angle falling in a specific 60° range, whereas no reinforcement was programmed for saccades in other directions.

Figure 6A shows data from one individual and Figure 6B from all six participants. During baseline on

←

the high-probability reinforcement schedules for all participants. These proportions were computed over baseline trials (for the first experimental conditions) or over the last 100 trials of the previous condition (for the second experimental conditions), and plotted against these proportions computed over the last 100 trials of each condition. Horizontal and vertical lines: chance proportions. Open symbols represent the first experimental condition, filled symbols the second experimental condition. Data from participant CO are indicated by circles.
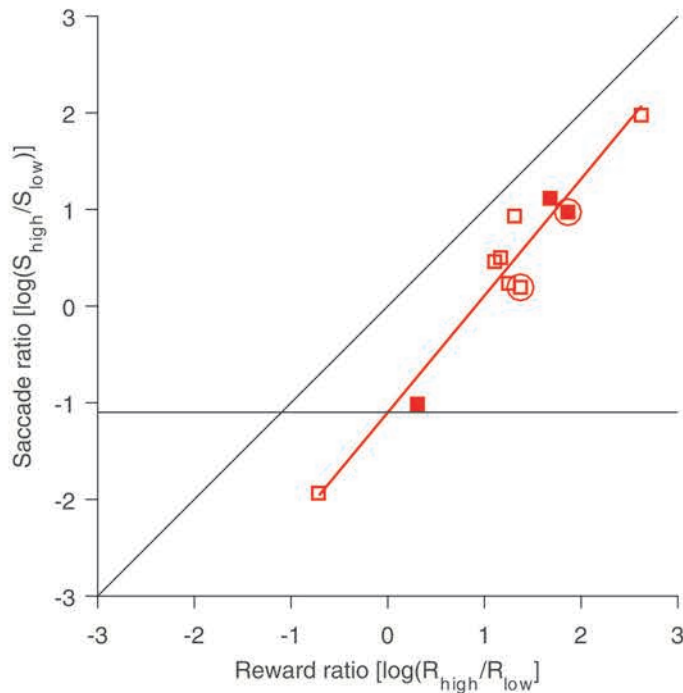
Figure 5. Matching analysis of the location experiment: log ratio of saccades meeting the high-probability reinforcement criteria and the low-probability reinforcement criteria plotted against the log ratio of number of targets seen under each alternative. Each data point represents data from one observer obtained over the last 100 trials of each condition. Filled data points are from the second condition of a participant with a different reinforcement schedule. Other conventions are the same as in Figure 3A.

average 14% ($SD = 4$) of observers' saccades were made in the frequently reinforced direction. This overall proportion reached 32% ($SD = 9$) at the end of the first experimental condition. The increase was significant for the whole group of participants, $t(5) = 6.51$, $p = 0.001$, as well as for all of the six individual participants, individual $\chi^2(1)$ tests, all $p < 0.001$. A similar increase from 9% (SD 13%) to 49% (SD 39%) was present if we considered only the first saccade in each trial, $t(5) = 2.75$, $p = 0.040$.

In this experiment, the proportion of saccades in the reinforced direction varied substantially between 22% and 46%, with reinforcement rates also varying between participants (Table 1 in Appendix). We performed the same data analysis using the generalized matching equation (Baum, 1974). We found that the data of experiment 3 were not very well described by the equation, when ratios of responses and reinforcers were calculated in terms of saccade direction (blue symbols in Figure 7; $R^2 = 0.18$). One potential reason for this weak correlation between saccadic directions and targets presented contingently on this property might be that participants associated the detection of a

target not with the direction of the saccade, but with a specific location.

We performed a second matching analysis to examine this possibility. We reanalyzed the data of this third experiment in terms of saccade landing positions using the same equation (red data in Figure 7). The quality of the model fit substantially increased ($R^2 = 0.95$) and the slope (0.81, 95% CI [0.60; 1.01]) was close to unity. As in the previous experiment, we observed a significant bias towards the areas of the search display associated with a low probability of seeing a target (-0.71; 95% CI [-0.87; -0.55]). Interestingly, reanalyzing the data of the position experiment in terms of directions resulted in a weak relationship ($R^2 = 0.30$) between saccade directions and reinforcement. These results suggest that it was indeed the location of the search target that was most relevant for successful learning in both experiments.

Discussion

Our results show that finding a search target at a particular location increases the probability that observers make saccades to that region in a future search. The act of finding the target, without any monetary or other reward associated with it, serves as a reward during visual search. Whereas it is neither possible nor desirable to control eye movements of observers while they perform the task, our gaze-contingent target presentation paradigm allowed us to fully control the reinforcement schedule. Our results provide direct evidence that complex eye movement behavior is sensitive to reinforcement by visual information gain, as hypothesized by earlier investigators (Clarke et al., 2016; Jiang et al., 2013).

The strength of our gaze-contingent paradigm lies in the control of the feedback function, namely, the relationship between responses (saccades) and their reinforcing consequences (finding the target). That is, we were able to determine the number of saccades participants had to make in one specific area or in one direction before detecting the target. This control differentiates our experiments not only from the studies on visual search conducted within the information theory framework mentioned previously (Najemnik & Geisler, 2005, 2008; Renninger et al., 2007), but also from studies on the emergence of attentional biases in space (Chun & Turk-Browne, 2008; Jiang et al., 2013; Jones & Kaschak, 2012; Kabata & Matsumoto, 2012; Peterson & Kramer, 2001; Walthew & Gilchrist, 2006). The latter investigated how people learn to prioritize important locations (those which are more likely to contain the target) from experience, based on regularities in the environment. Typically target location

Figure 7. Matching analysis of the direction experiment. Data from participant FL are indicated by circles. Blue indicates data analyzed in terms of saccade direction; red indicates data analyzed in terms of saccade landing position. Other conventions are the same as in Figure 5.

probability is manipulated to be higher at a particular location than at the others—independently from the observer's eye movements. This manipulation induces search time benefits for targets in high-probability locations and more frequent first saccades directed towards these locations. Along the same lines, eye movement strategies in a face discrimination task could be modified by shifting the regions containing maximum information for solving the task. For regular faces, the eye region is most relevant for discrimination, and observers tend to fixate there more than in other regions. When the stimuli differed mainly in the mouth region, participants learned to directly orient their gaze towards the specific task-relevant region (Peterson & Eckstein, 2014). It would be worthwhile to use our gaze-contingent paradigm first to examine whether the direct manipulation of reward led to an optimal search strategy and, second, to compare the learning rate with an ideal observer that takes into account the statistics of the task.

←

to 55% during the first experimental condition and from 4% to 32% during the second condition. (B) Global results. Proportions of saccades meeting the criteria of the high-probability reinforcement schedules for all participants. Conventions are the same as in Figure 4.
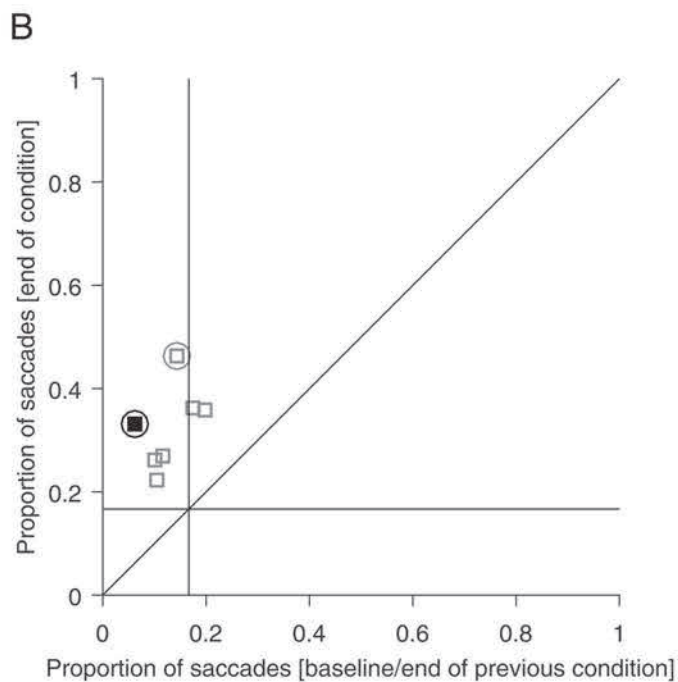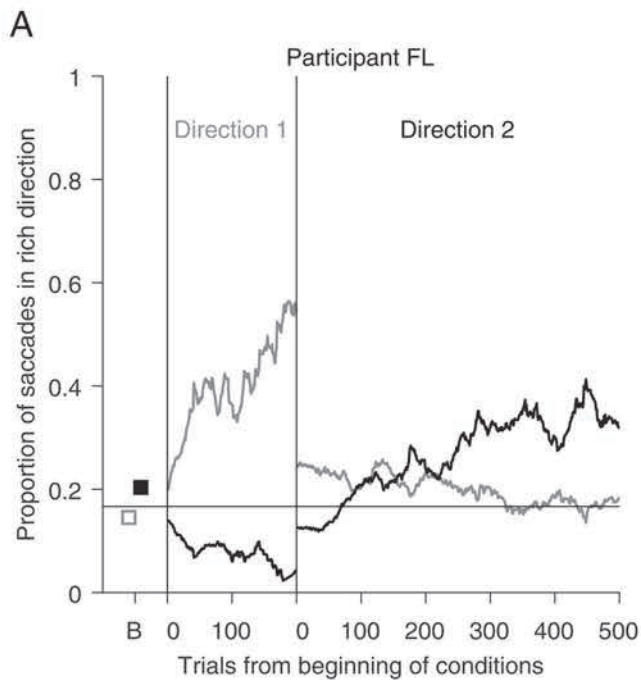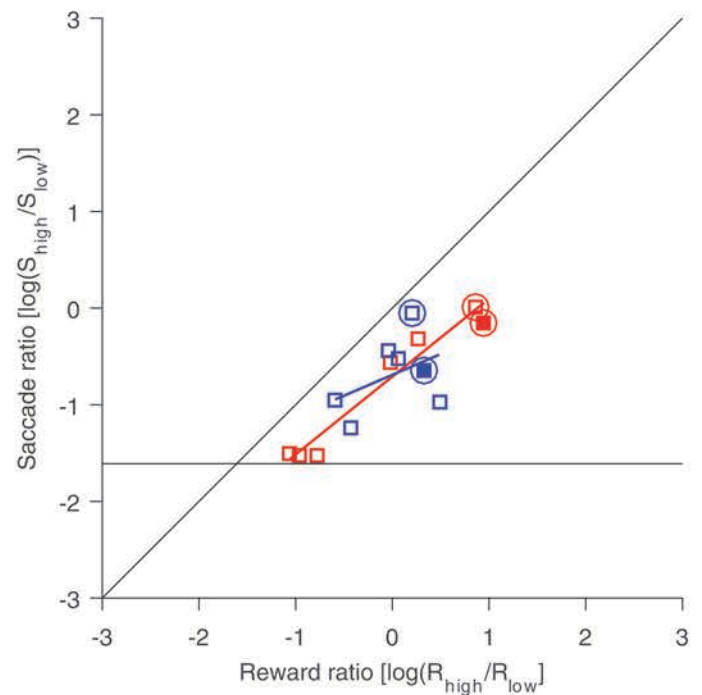

Figure 6. Results of the direction experiment. (A) Evolution of the proportion of saccades meeting the criteria of the high-probability reinforcement schedules for participant FL. During the first daily session we programmed to reinforce 100% of the saccades located within a specific 60° range of angles. During the second session, the reinforcement rate was cut in half and then remained constant until the end of the experiment, even during the second condition following the change of reinforced angle range. No reinforcement for other saccades was programmed in this experiment. For this participant, the proportion of saccades moving at an angle contained in the reinforced range increased following our procedure, from 14%

It should be noted that our gaze-contingent paradigms of experiments 2 and 3 and the manipulation of the target location probability in experiment 1 also differed with respect to the sensory evidence present when the search is initiated. In the latter case, a task-related visual signal was present from the beginning of the trials. In the reinforcement paradigms, there was only sensory noise. In statistical decision theory, reward and prior knowledge about probability distributions associated with world states are central to understanding motor control, but uncertainty also plays a major role (e.g., Glimcher, 2003; Shadmehr, 2009; Trommershäuser, Glimcher, & Gegenfurtner, 2009; Trommershäuser, Maloney, & Landy, 2008). In visual search paradigms, it has been shown that observers optimally combine prior spatial beliefs about where the target may occur with current sensory evidence (Eckstein, Drescher, & Shimozaki, 2006; Torralba, Oliva, Castelhano, & Henderson, 2006; Vincent, 2011). Similarly, participants might have attributed more weight to top-down information in our reinforcement paradigms than in the frequency experiment. This difference might have contributed to the discrepancy observed in the results. In the future, the elaboration of a dynamic model of visual search learning should also take into account the influence of visual factors such as target contrast and the visibility of the target at different eccentricities. However, this is beyond the scope of the present study, and further experiments are necessary to manipulate these factors systematically.

In contrast to these earlier experiments with a small limited number of target locations, we did not find a modification in eye movement strategy when we only changed biases in the occurrence of stimuli in different regions. The major difference to these experiments is that the search target could occur at a very large number of possible locations in our case. In the aforementioned studies on attentional biases, the visual system has to take into account only a small number of spatially distinct alternatives. Even though performance in forced-choice paradigms involving few locations can be used to predict results obtained with many locations (Burgess & Ghandeharian, 1984), these tasks are not equivalent to free localization tasks which, for instance, do not require the maintenance of a detection criterion from trial to trial (Abbey & Eckstein, 2014). The differing outcome in our case additionally emphasizes that the results of such paradigms cannot be readily transferred to realistic search conditions with many potential target locations. In this case it would simply take a very long time to acquire a statistical bias. By providing reinforcement contingently on saccadic eye movements, we were able to quickly induce modifications in search strategies, thereby demonstrating the control of saccadic eye movements by their own visual consequences—even in complex search situations.

In the brain, the coding of saccades does not follow world coordinates, but they are determined by retinal direction and amplitude (Lee, Rohrer, & Sparks, 1988; Wurtz & Albano, 1980). Due to the geometry of our search display, angular position and movement direction of saccades were correlated in experiment 2, circular correlation coefficient, $r(45322) = 0.49$, $p < 0.001$. It might have been possible that the learning of saccadic consequences occurred for direction and amplitude, rather than for position on the screen. However, the matching analyses showed that participants matched the relative frequency of their saccades made towards a specific area of the screen to the relative number of targets seen in these areas. This was true regardless of whether this visual consequence was contingent on landing positions (experiment 2, Figure 5) or on saccade directions (experiment 3, Figure 7). Presumably, compared to the relation between target detection and saccadic direction, the relation between target location and gaze allocation during visual search reflects a more natural contingency that observers have learned for years.

Despite the observed matching behavior, the participants did not favor the highly reinforced areas of the search display as much as they should have, based on the experienced reinforcement. This is shown by the nonzero intercept of the matching equation fit in Figures 5 and 7. Several hypotheses can be put forward to explain this bias in favor of the areas associated with a lower probability of reinforcement. First, in our paradigm these regions represent a much larger area than the rich quadrant, and there might be a natural tendency for eye movements to distribute within the whole search region. Second, perceived reinforcement could be weaker than the scheduled reinforcement, because we cannot always be certain to which saccade the observers attributed target detection. Third, participants might prefer exploration rather than maximum exploitation. The exploration-exploitation trade-off is a concept often used in the reinforcement learning approach of machine learning (Sutton & Barto, 1998). Exploration offers an advantage in changing environments because it allows detecting new sources of reinforcement for further exploitation. In our experiment, participants might check if a target could be present in the low reinforcement quadrants in order to avoid missing it or to confirm that the reinforcement rates did not change.

The fact that human observers match relative saccade rates to relative visual reinforcement rates is consistent with a vast number of behavioral studies on the matching law, which describes how individuals tend to respond in proportion to the expected value of behavioral consequences (Herrnstein, 1961; Poling,

Edwards, Weeden, & Foster, 2011). These studies were conducted in many different species with animals activating keys or levers associated with various reinforcement rates, in order to investigate how different organisms optimize their decisions in probabilistic environments. Our study is the first one reporting matching behavior in a human sensorimotor system in the context of an unconstrained visual search task.

In primates, a linear relationship between the frequency of saccades made towards specific targets and the relative probability of alimentary reward signaled by these targets has been found (Sugrue, Corrado, & Newsome, 2004). In such neurophysiological studies eye movements were investigated, but mainly used as a way to express binary decisions. Contrary to our visual search paradigm, these choices were not spatially extended, and were merely used to investigate neural processes underlying target selection. Indeed, neural activity changes correlating with alimentary reward have been observed in the basal ganglia (Hikosaka et al., 2006), specifically in the caudate nucleus (Lauwereyns, Watanabe, Coe, & Hikosaka, 2002) and the substantia nigra pars reticulate (Sato & Hikosaka, 2002), as well as the superior colliculus (Ikeda & Hikosaka, 2003). It has also been found that dopaminergic neurons in the brainstem and the frontal eye field (Soltani, Noudoost, & Moore, 2013) contribute to saccadic target selection. Moreover a growing body of evidence suggests that cholinergic and serotoninergic brainstem neurons might encode the predicted and actual reward value associated with saccadic targets (see Okada, Nakamura, & Kobayashi, 2011, for a review). According to these studies, the reward system could constitute a neural basis for learning eye movement behavior. This system might also be involved in learning a complex visual search task.

Conclusion

Visual search, and human actions in general, have been frequently characterized as being optimal (e.g., Eckstein et al., 2015; Najemnik & Geisler, 2005, 2008; Wolpert & Ghahramani, 2000) or at least near optimal (Ackermann & Landy, 2013; Droll et al., 2009). It has been unclear how observers would acquire such optimality, faced with the interactions of many complex constraints, like their own motor variability or differences in target visibility. Our results show that finding a visual target can serve as a direct reward, and reinforcement learning can serve as the mechanism to optimize eye movement strategies. This way, reinforcement learning could form the basis of intelligent

behavior without requiring prior knowledge or extensive computations. Using our new gaze-contingent reward paradigm, we could show why we look more often in specific locations when we search for particular items. Researchers being able to modify search behavior rapidly also has the potential to be useful in training novice searchers in new and unfamiliar environments.

Keywords: reinforcement learning, eye movements, visual search, matching analysis

References

Abbey, C. K., & Eckstein, M. P. (2014). Observer efficiency in free-localization tasks with correlated noise. *Frontiers in Psychology*, 5, 1–15.

Ackermann, J. F., & Landy, M. S. (2013). Choice of saccade endpoint under risk. *Journal of Vision*, *13*(3):27, 1–20, doi:10.1167/13.3.27. [PubMed] [Article]

Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences*, 24, 191–204.

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, *22*(1), 231–242.

Berlyne, D. E. (1972). Reinforcement values of visual patterns compared through concurrent performances. *Journal of the Experimental Analysis of Behavior*, *18*(2), 281–285.

Burgess, A. E., & Ghandeharian, H. (1984). Visual signal detection. II. Signal-location identification. *Journal of the Optical Society of America A*, *1*(8), 906–910.

Chen, X., Mihalas, S., Niebur, E., & Stuphorn, V. (2013). Mechanisms underlying the influence of

saliency on value-based decisions. *Journal of Vision*, *13*(12):18, 1–23, doi:10.1167/13.12.18. [PubMed] [Article]

Chukoskie, L., Snider, J., Mozer, M. C., Krauzlis, R. J., & Sejnowski, T. J. (2013). Learning where to look for a hidden target. *Proceedings of the National Academy of Sciences, USA, 110* (Suppl. 2), 10438–10445.

Chun, M. M., & Turk-Browne, N. B. (2008). Associative learning mechanisms in vision. In S. J. Luck & A. Hollingsworth (Eds.), *Visual memory (Oxford series in visual cognition)* (pp. 209–245). Oxford, UK: Oxford University Press.

Clarke, A. D., Green, P., Chantler, M. J., & Hunt, A. R. (2016). Human search for a target on a textured background is consistent with a stochastic model. *Journal of Vision*, *16*(7):4, 1–16, doi:10.1167/16.7.4. [PubMed] [Article]

Collins, T. (2012). Probability of seeing increases saccadic readiness. *PLoS One*, *7*(11), e49454.

Dorris, M. C., Pare, M., & Munoz, D. P. (2000). Immediate neural plasticity shapes motor performance. *Journal of Neuroscience*, *20*(1), 1–5.

Droll, J. A., Abbey, C. K., & Eckstein, M. P. (2009). Learning cue validity through performance feedback. *Journal of Vision*, *9*(2):18, 1–22, doi:10.1167/9.2.18. [PubMed] [Article]

Eckstein, M. P. (2011). Visual Search: A retrospective. *Journal of Vision*, *11*(5):14, 1–36, doi:10.1167/11.5.14. [PubMed] [Article]

Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science*, *17*(11), 973–980.

Eckstein, M. P., Schoonveld, W., Zhang, S., Mack, S. C., & Akbas, E. (2015). Optimal and human eye movements to clustered low value cues to increase decision rewards during search. *Vision Research*, *113*, 137–154.

Glimcher, P. W. (2003). The neurobiology of visual-saccadic decision making. *Annual Review of Neuroscience*, *26*, 133–179.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*, 267–272.

Hikosaka, O., Nakamura, K., & Nakahara, H. (2006). Basal ganglia orient eyes to reward. *Journal of Neurophysiology*, *95*(2), 567–584.

Ikeda, T., & Hikosaka, O. (2003). Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron*, *39*(4), 693–700.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506.

Jiang, Y. V., Swallow, K. M., Rosenbaum, G. M., & Herzig, C. (2013). Rapid acquisition but slow extinction of an attentional bias in space. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 87–99.

Jones, J. L., & Kaschak, M. P. (2012). Global statistical learning in a visual search task. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 152–160.

Kabata, T., & Matsumoto, E. (2012). Cueing effects of target location probability and repetition. *Vision Research*, *73*, 23–29.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, 14.

Lauwereyns, J., Watanabe, K., Coe, B., & Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature*, *418*(6896), 413–417.

Lee, C., Rohrer, W. H., & Sparks, D. L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature*, *332*(24), 357–360.

Madelain, L., Paeye, C., & Wallman, J. (2011). Modification of saccadic gain by reinforcement. *Journal of Neurophysiology*, *106*(1), 219–232.

Montagnini, A., & Chelazzi, L. (2005). The urgency to look: Prompt saccades to the benefit of perception. *Vision Research*, *45*(27), 3391–3401.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*(7010), 760–767.

Morvan, C., & Maloney, L. T. (2012). Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS One*, *8*(2), e1001342.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387–391.

Najemnik, J., & Geisler, W. S. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision*, *8*(3):4, 1–14, doi:10.1167/8.3.4. [PubMed] [Article]

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research*, *51*(13), 1526–1537.

Navalpakkam, V., Koch, C., Rangel, A., & Perona, P. (2010). Optimal reward harvesting in complex perceptual environments. *Proceedings of the National Academy of Sciences, USA*, *107*(11), 5232–5237.

Okada, K., Nakamura, K., & Kobayashi, Y. (2011). A neural correlate of predicted and actual reward

value information in monkey pedunculopontine tegmental and dorsal raphe nucleus during saccade tasks. *Neural Plasticity, 2011*, 1–21.

Paeye, C., & Madelain, L. (2011). Reinforcing saccadic amplitude variability. *Journal of the Experimental Analysis of Behavior, 95*(2), 149–162.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research, 42*(1), 107–123.

Peterson, M., & Eckstein, M. P. (2014). Learning optimal eye movements to unusual faces. *Vision Research, 99*, 57–68.

Peterson, M., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics, 63*(7), 1239–1249.

Poling, A., Edwards, T. L., Weeden, M., & Foster, T. M. (2011). The matching law. *The Psychological Record*, (61), 313–322.

Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements reduce local uncertainty. *Journal of Vision, 7*(3):6, 1–17, doi:10.1167/7.3.6. [PubMed] [Article]

Sato, M., & Hikosaka, O. (2002). Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement. *Journal of Neuroscience, 22*(6), 2363–2373.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision, 11*(5):9, 1–30, doi:10.1167/11.5.9. [PubMed] [Article]

Schütz, A. C., Kerzel, D., & Souto, D. (2014). Saccadic adaptation induced by a perceptual task. *Journal of Vision, 14*(5), 4, 1–19, doi:10.1167/14.5.4. [PubMed] [Article]

Shadmehr, R. (2009). Computational approaches to motor control. In L. R. Squire (Ed.), *Encyclopedia of neuroscience* (Vol. 3, pp. 9–17). San Diego, CA: Elsevier.

Soltani, A., Noudoost, B., & Moore, T. (2013). Dissociable dopaminergic control of saccadic target selection and its implications for reward modulation. *Proceedings of the National Academy of Sciences, USA, 110*(9), 3579–3584, doi.org/10.1073/pnas.1221236110.

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science, 304*(5678), 1782–1787.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT press.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision, 11*(5):5, 1–23, doi:10.1167/11.5.5. [PubMed] [Article]

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766–789.

Towal, R. B., Mormann, M., & Koch, C. (2013). Simultaneous modeling of visual saliency and value computation improves predictions of economic choice. *Proceedings of the National Academy of Sciences, USA, 110*(40), E3858–E3867.

Trommershäuser, J., Glimcher, P. W., & Gegenfurtner, K. R. (2009). Visual processing, learning and feedback in the primate eye movement system. *Trends in Neurosciences, 32*(11), 583–590.

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2008). Decision making, movement planning and statistical decision theory. *Trends in Cognitive Sciences, 12*(8), 291–297.

Verghese, P. (2010). Active search for multiple targets is inefficient. *Journal of Vision, 10*(7): 1296, doi:10.1167/10.7.1296. [Abstract]

Vincent, B. (2011). Covert visual search: Prior beliefs are optimally combined with sensory evidence. *Journal of Vision, 11*(13):25, 1–15, doi:10.1167/11.13.25. [PubMed] [Article]

Walthew, C., & Gilchrist, I. D. (2006). Target location probability effects in visual search: An effect of sequential dependencies. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1294–1301.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113–20.

Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search, In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York, NY: Oxford.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience, 3* (Suppl.), 1212–1217.

Wurtz, R. H., & Albano, J. E. (1980). Visual-motor function of the primate superior colliculus. *Annual Review of Neuroscience, 3*(1), 189–226.

Xu-Wilson, M., Zee, D. S., & Shadmehr, R. (2009). The intrinsic value of visual information affects saccade velocities. *Experimental Brain Research, 196*(4), 475–481.

Appendix

Participant	Nb of sess	Condition 1: Learned quadrant				Condition 2: Relearned quadrant			
		Number of trials	Mean nb sac per trial	Program reinft ratio	Obtained reinft ratio	Number of trials	Mean nb sac per trial	Program reinft ratio	Obtained reinft ratio
Experiment 1									
KI	10	1363	3.5		0.69 : 0.36				
MK	5	849	6.7		0.23 : 0.21				
JU	5	1329	5.8		0.23 : 0.15				
AH	3	954	6.3		0.30 : 0.21				
MI	4	773	9.1		0.22 : 0.13				
SO	5	955	5.6		0.27 : 0.15				
Experiment 2									
MA	4	584	18.5	0.4 : 0.1	0.19 : 0.06				
RO	5	991	7.3	0.8 : 0.2	0.36 : 0.13				
FR	4	296	4.4	0.8 : 0.2	0.33 : 0.17	494	8.1	0.4 : 0.1	0.18 : 0.10
AK	4	1089	6.4	0.4 : 0.1	0.15 : 0.10				
CO	3	495	7.4	0.4 : 0.1	0.23 : 0.07	299	8.4	0.4 : 0.1	0.17 : 0.07
CA	4	892	11.4	0.4 : 0.1	0.15 : 0.08				
MO	5	497	9.9	0.4 : 0.1	0.16 : 0.09	691	10	0.4 : 0.1	0.21 : 0.06
Experiment 3									
DE	5	1167	13.7	1 : 0	0.19 : 0.11				
FL	4	199	5.7	1 : 0	0.33 : 0.14	496	11.6	0.5 : 0	0.18 : 0.06
JA	1	288	6.4	1 : 0	0.32 : 0.18				
AN	4	885	10.2	0.5 : 0	0.15 : 0.09				
RB	8	955	26.7	0.5 : 0	0.09 : 0.06				
VP	8	1380	15.9	0.5 : 0	0.16 : 0.07				

Table 1. Total number of trials, mean number of saccades per trial, programmed and obtained ratios of reinforcement per condition, for each participant of each experiments. *Notes*: The first and second numbers in the reinforcement ratio columns refer to the high-probability and low-probability reinforcement schedules, respectively. Nb of sess: Number of one-hour daily sessions; Nb sac: Number of saccades; Program reinft ratio: Programmed reinforcement ratio; Obtained reinft ratio: Ratio of reinforcement actually obtained by the participants.