



# Reward learning biases the direction of saccades

Ming-Ray Liao\*, Brian A. Anderson

Texas A&M University, United States of America



## ARTICLE INFO

### Keywords:

Eye movements  
Overt attention  
Reward learning  
Reinforcement  
Visual search  
Foraging

## ABSTRACT

The role of associative reward learning in guiding feature-based attention and spatial attention is well established. However, no studies have looked at the extent to which reward learning can modulate the direction of saccades during visual search. Here, we introduced a novel reward learning paradigm to examine whether reward-associated directions of eye movements can modulate performance in different visual search tasks. Participants had to fixate a peripheral target before fixating one of four disks that subsequently appeared in each cardinal position. This was followed by reward feedback contingent upon the direction chosen, where one direction consistently yielded a high reward. Thus, reward was tied to the direction of saccades rather than the absolute location of the stimulus fixated. Participants selected the target in the high-value direction on the majority of trials, demonstrating robust learning of the task contingencies. In an untimed visual foraging task that followed, which was performed in extinction, initial saccades were reliably biased in the previously rewarded-associated direction. In a second experiment, following the same training procedure, eye movements in the previously high-value direction were facilitated in a saccade-to-target task. Our findings suggest that rewarding directional eye movements biases oculomotor search patterns in a manner that is robust to extinction and generalizes across stimuli and task.

## 1. Introduction

Our perceptual systems have a limited representational capacity, making selective attention important for survival. The influence of associative reward learning and reinforcement learning in stimulus-evoked attentional orienting has been well established. Previously reward-predictive features capture attention in a subsequent visual search task in which those features are nonsalient and task-irrelevant (see Anderson, 2016, for a recent review). Value-driven attentional capture has been successfully demonstrated using a variety of stimulus features, including color (Anderson, Laurent, & Yantis, 2011), orientation (Laurent, Hall, Anderson, & Yantis, 2015; Theeuwes & Belopolsky, 2012), shape (Della Libera & Chelazzi, 2009; Della Libera, Perlato, & Chelazzi, 2011), and object category (Donohue et al., 2016; Hickey, Kaiser, & Peelen, 2015).

Recently, Anderson and Kim (2018a, 2018b) extended the principle of value-driven attention to the orienting of attention towards reward-associated spatial locations within a scene (see also, Chelazzi et al., 2014). Participants performed an initial task where an attentional shift to a particular region of space resulted in a reward. This was followed by either a visual search task superimposed on the scenes experienced

during training, or a free viewing task in which participants were exposed to these same scenes. In the visual search task, targets were identified faster when appearing in the previously reward-associated region if the search stimuli appeared soon after scene onset, while inhibition of return was evident at a longer stimulus onset asynchrony, reflecting a scene-evoked spatial orienting response. In the free viewing tasks, previously reward-associated regions were associated with both more frequent saccades in their direction and longer fixation durations, and this bias was evident from the initial saccade. These authors further showed that the bias is not contingent upon the richness of the scene, in that the bias can occur in the absence of objects in a free viewing task (Anderson & Kim, 2018b). However, this same bias was not evident using object-less scenes during goal-directed visual search (Anderson & Kim, 2018a), which could reflect the ability to override value-driven spatial attentional biases when performing an explicit task.

Although the influence of reward learning on visual information processing has been investigated for spatial locations, as measured by fixations and saccades (Anderson & Kim, 2018a, 2018b), the influence of reward learning on the oculomotor system itself has not been investigated. Saccades are rapid, ballistic movements of the eye that change the point of fixation, which can be triggered reflexively (Purves,

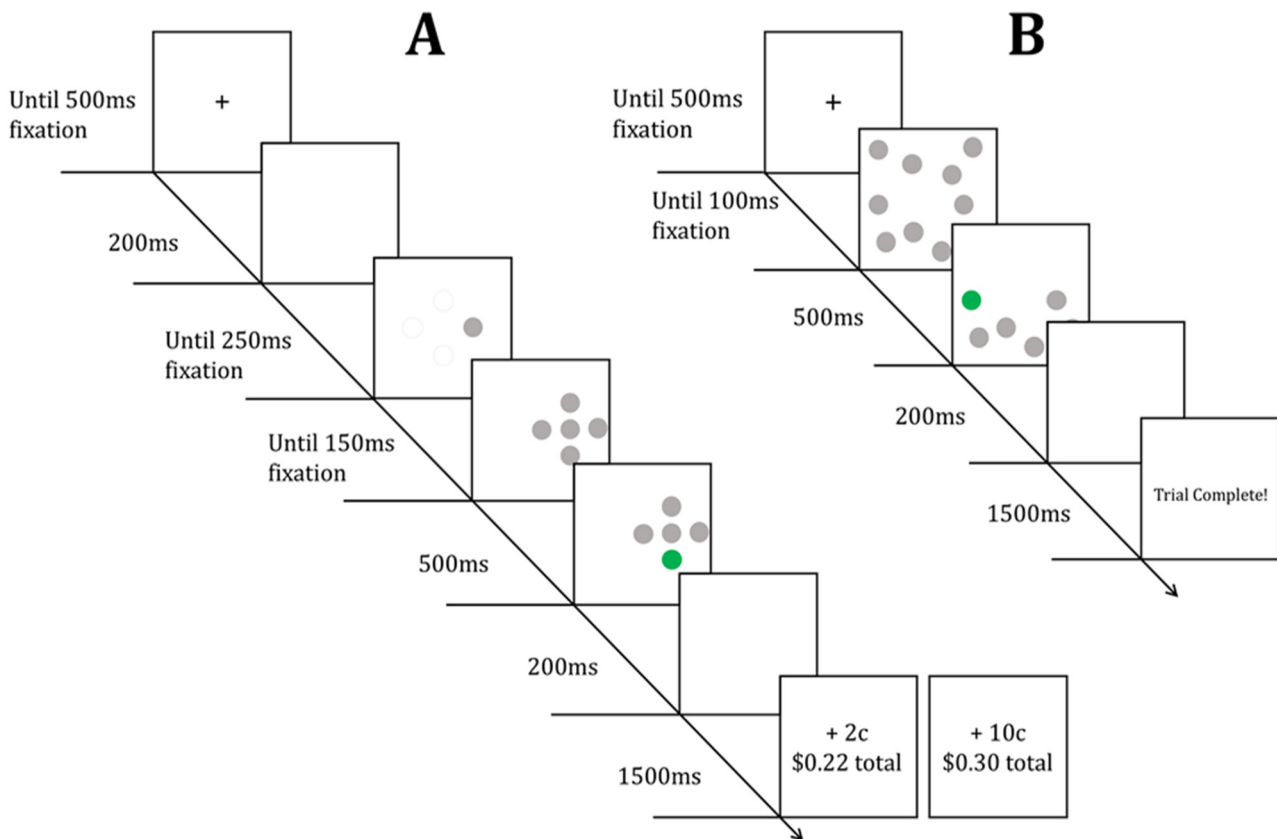
\* Corresponding author at: Texas A&M University, Department of Psychological & Brain Sciences, 4235 TAMU, College Station, TX 77843-4235, United States of America.

E-mail address: [m4liao@tamu.edu](mailto:m4liao@tamu.edu) (M.-R. Liao).

<https://doi.org/10.1016/j.cognition.2019.104145>

Received 19 July 2019; Received in revised form 14 October 2019; Accepted 16 November 2019

0010-0277/ © 2019 Elsevier B.V. All rights reserved.



**Fig. 1.** Time course of trial events during the training and test phase of Experiment 1. (A) Participants fixated on the first circle before choosing one of four circles to fixate on. Note that the dotted-line circles indicate the other locations where the first circle could also appear and are for illustration purposes only (i.e., did not appear in the actual task). The reward was contingent upon the direction of their choice. In this example, the bottom direction is the high-value direction, and fixating on the bottom target yielded the highest possible reward. The same procedure was used for Experiment 2 albeit with a minor modification – the first circle fixated was a fixation cross instead of another circle. (B) Participants fixated on the cross before choosing a circle by fixating on it. The chosen circle will disappear or turn green to indicate that the target has been found. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Augustine, & Fitzpatrick, 2012); this allows us to foveate stimuli of potentially high interest, maximizing our ability to process these stimuli with high visual acuity. Neuroeconomic studies in monkeys have shown that reward seems to have a direct influence on saccade programming (Bendiksby & Platt, 2006; Ikeda & Hikosaka, 2003; Kawagoe, Takikawa, & Hikosaka, 1998; Milstein & Dorris, 2011; Takikawa, Kawagoe, & Hikosaka, 2002; Yamamoto, Kim, & Hikosaka, 2013). For example, Sohn and Lee (2006) showed that even though a target remained on the display, the animals would consistently generate initial errant saccades to a task-irrelevant location in space previously associated with reward. In this situation, the reward-associated region is devoid of stimuli but still able to elicit saccades.

Prior studies have also shown that reward contingencies can speed up saccades (Bendiksby & Platt, 2006; Kawagoe et al., 1998; Takikawa et al., 2002; Yamamoto et al., 2013). Similar results have been obtained in studies using human observers (Bucker, Silvis, Donk, & Theeuwes, 2015; Milstein & Dorris, 2007; Theeuwes & Belopolsky, 2012). For example, in Chen and colleagues' study, participants were cued to make saccades to the left or right and were also cued to whether or not the trial would be a reward trial. On potential reward trials, participants showed increased peak velocity in saccades (Chen, Chen, Zhou, & Mustain, 2014).

Directional saccades reflect an overt behavior, which might be subject to biases arising from reinforcement learning. However, it remains unclear whether reward learning can bias the generation of eye movements in a particular direction, separately from the particular region of space to which they are made. That is, it is unclear if reward learning can bias the direction of eye movements per se. In the present

study, we associated reward with saccades made in a particular cardinal direction in a spatial orienting task. Participants were provided monetary incentives encouraging them to orient towards a specific cardinal direction, which we refer to as the “high-value” direction.

The high-value direction was decoupled from specific regions of space by having participants fixate initially at the center of the screen, and then to one of four cardinal directions before a decision-making task occurred. The decision-making task consisted of saccading to one of four peripheral targets, placed with random jitter in each of the four cardinal directions from the point of current fixation. Saccading to a target in one direction was associated with more reward (probabilistically) than saccading to a target in any of the other directions. In this task, the initial saccade from central fixation is, in principle, analogous to the different contexts (scenes) in Anderson and Kim's (2018a, 2018b) paradigm, as it separated the rewarded direction from a specific spatial location on the screen. We examined whether participants could robustly learn which direction was the optimal direction in which to orient during training, and whether such learning would transfer to performance in an unrewarded visual search task. To this latter end, we measured eye movements after training in both an unguided (visual foraging, Experiment 1) and guided (saccade to a shape-defined target, Experiment 2) visual search tasks, given the potential distinction between the two as evident in value-driven attention to spatial locations (see Anderson & Kim, 2018a, 2018b).

## 2. Experiment 1

### 2.1. Methods

#### 2.1.1. Participants

Forty participants (18–35 years of age,  $M = 22.2$  years; 25 females, 15 males) were recruited from the Texas A&M Community. Participants were compensated with money earned in the experimental task. All reported normal or corrected-to-normal visual acuity and normal color vision. The data from one participant were dropped and replaced with data from a new participant due to an inability to reliably track eye position (resulting in a failure to register a target fixation on over 30% of trials); thus, 41 individuals in total were consented and participated. All procedures were approved by the Texas A&M University Institutional Review Board and conformed with the principles outlined in the Declaration of Helsinki. The sample size of  $n = 40$  was determined a priori and would yield  $\beta > 0.80$  to detect an effect as small as  $d_z = 0.46$  (computed using G\*Power 3.1), which was smaller than the effect size for reward learning on eye movements evident in Anderson and Kim (2018a, 2018b).

#### 2.1.2. Apparatus

A Dell OptiPlex equipped with Matlab software and Psychophysics Toolbox extensions (Brainard, 1997) was used to present the stimuli on a Dell P2717H monitor. The participants viewed the monitor from a distance of approximately 70 cm in a dimly lit room. Eye position was monitored using an EyeLink 1000-plus desktop-mount eyetracker (SR Research, Ottawa, Ontario, Canada). Head position was maintained using an adjustable chin rest (SR Research).

#### 2.1.3. Training phase

Each trial consisted of a fixation display, a secondary fixation display, a choice array, and a reward feedback display (see Fig. 1A). The fixation cross ( $1.3^\circ$  visual angle) remained on the screen until eye position had been registered within  $1.8^\circ$  of the fixation cross for a continuous period of 500 ms. After a 200 ms blank screen, a circle would then appear in one of four cardinal directions at  $5.8^\circ$  eccentricity and remain until eye position had been registered within  $3.4^\circ$  of the center of the circle for a continuous period of 250 ms. A choice array consisting of four grey circles jittered (randomly up to  $20^\circ$  of arc in either direction from the cardinal positions) on an imaginary circle with a radius of  $8.2^\circ$  was then presented until fixation on a target was registered for a continuous period of 150 ms. The grey circles were approximately  $1.0^\circ$  in radius, and a region extending  $2.5^\circ$  beyond the boundary of the circle was used as an area of interest for determining target fixations. Once a fixation on a target was registered, the chosen circle would turn green for 500 ms before a 200 ms blank screen. After which, the reward feedback display was presented for 1500 ms, and consisted of the money earned on the current trial along with the updated total earnings. Each trial concluded with a 1500 ms blank interval.

Participants were instructed to fixate (“look directly at”) the cross to begin each trial, then to fixate on the first circle, and then choose the best one out of four circles to look at. The first circle appeared randomly in the four cardinal directions, with the constraint that a circle appeared in each position equally-often in each block of the task. Fixating on a circle appearing in one direction (up, right, down, or left, counterbalanced across participants) in the display with four peripheral circles was associated with an 80% probability of a high reward of 10¢, and a 20% probability of a low reward of 2¢, while for the other directions these percentages were reversed. Each block consisted of 60 trials, the order of which was randomized.

#### 2.1.4. Test phase

Each trial consisted of a fixation display (until fixation was acquired for a continuous period of 500 ms), a search array until target was

found, and a 1500 ms feedback screen (see Fig. 1B). The search array consisted of 10 circles (two in two quadrants and three in two quadrants, counterbalanced across trials), with each circle  $0.9^\circ$  in radius. On each trial the circles were at least  $2.9^\circ$  away from the edge of the screen,  $6.4^\circ$  away from any other circle and  $7.4^\circ$  away from the center of the screen.

A region extending  $2.5^\circ$  beyond the boundary of the circle was used as an area of interest for determining stimulus fixations. Participants were instructed to look directly at different circles until the target was found. There was no explicit time pressure, and participants could not time out. Selection of a circle was registered after 100 ms of continuous fixation on that circle, after which the circle would disappear if it was not the target or turn green for 500 ms to indicate that the target had been found. The feedback display consisted of the words “Trial Complete!” presented at the center of the screen. The target appeared equally-often in each quadrant. Each block consisted of 48 trials (with each combination of 3 and 2 stimuli per quadrant presented eight times), the order of which was randomized.

#### 2.1.5. Procedure

Participants completed 4 blocks of trials of the training phase, followed by 3 blocks of trials of the test phase. Both the training and test phases were preceded by interactive instructions that included practice trials (4 for the training phase and 6 for the test phase). Participants were paid the amount of money earned in the training phase at the completion of the experiment (mean = \$16.81, range = \$8.96–\$20.08).

#### 2.1.6. Measurement of eye position

Head position was maintained throughout the experiment using an adjustable chin rest that included a bar upon which to rest the forehead (SR Research). Participants were provided a short break between different runs of the task, during which they were allowed to reposition their head to maintain comfort. Eye position was calibrated prior to each block of trials using a 9-point calibration (Anderson & Yantis, 2012) and was manually drift-corrected by the experimenter as necessary (the next trial could not begin until eye position had been registered within  $1.8^\circ$  visual angle of the center of the fixation cross for 500 ms; see, e.g., Nissens, Failing, & Theeuwes, 2016). During the presentation of the search array, the X and Y position of the eyes was continuously monitored in real time with respect to the six stimulus positions, such that fixations were coded online (Le Pelley, Pearson, Griffiths, & Beesley, 2015).

#### 2.1.7. Analysis of saccades

We calculated the direction of saccade in the test phase via the coordinates of the circle fixated. If the distance from the prior fixation to the next circle fixated was greater in the x than y dimension, then the direction of saccade was either to the left or right as opposed to up or down. The direction of saccades was summed up, separately for each saccade made in the trial (first, second, third, etc.), and the proportion of high-value choices was calculated. The resulting proportions were then corrected for training-independent biases to look in a particular direction by subtracting the proportion of choices in that high-value direction computed across participants in the other training conditions, and tested against zero using a one-sample *t*-test.<sup>1</sup> A parallel analysis was performed on the number of saccades required to find the target by the quadrant within which the target was located.

<sup>1</sup> The results remain significant if no correction for training-independent bias is applied.

## 2.2. Results

### 2.2.1. Training phase

Participants were able to learn the reward association, with 70.5% of all saccades going in the high-value direction (as opposed to 25% if participants made saccades in random directions),  $t(39) = 11.765$ ,  $p < 0.001$ ,  $d = 1.860$ . Saccades made in the high-value direction increased over the course of the task, from 52.7% in the first block to 81.3% in the last block,  $t(39) = 7.419$ ,  $p < 0.001$ ,  $d = 1.173$ . To determine whether the learning was robust to the direction of the initial (first-step) saccade, as would be predicted from a directional bias rather than a bias to orient to a particular region of space, we ran a paired samples t-test comparing the percentage of high-value target choice between first-step saccades in the rewarded direction and in the opposite direction and found no difference between them in any block ( $ps > 0.120$ ).

### 2.2.2. Test phase

Rewarding directional saccades during training produced an 8.1% increase in the frequency of initial saccades made in the high-value direction, which was significant,  $t(39) = 2.99$ ,  $p = 0.005$ ,  $d = 0.472$  (visual depiction of this effect is shown in Figs. 2 and 3). The magnitude of this bias did not significantly differ across the four training conditions,  $F(3,36) = 0.82$ ,  $p = 0.491$ . This bias was also robust to extinction, being individually significant in the first (7.9%) and last (6.9%) block of the test phase,  $ts > 2.49$ ,  $ps < 0.018$ ,  $ds > 0.39$ , with the difference between blocks being non-significant,  $t(39) = 0.45$ ,  $p = 0.658$ . For completeness, we also investigated the second and third saccades relative to the center and to the prior circle fixated. Only the second saccade relative to the first circle fixated was significant,  $t(39) = -2.39$ ,  $p = 0.022$ ,  $d = -0.379$  (other  $ps > 0.17$ ), reflecting a bias to look in a different direction.

Analysis of the number of saccades required to find the target mirrored the bias in the direction of the initial saccade, with fewer overall saccades observed when the target was in the high-value quadrant (mean = 5.19) compared to a low-value quadrant (mean = 5.50),  $t(39) = 2.20$ ,  $p = 0.034$ ,  $d = 0.35$ .

### 2.3. Discussion

Previous studies have investigated direction and speed of saccades towards a previously reward-associated object as an index of value-driven attentional capture (A.J. Kim and Anderson, 2019; Anderson &

Yantis, 2012; Bucker et al., 2015; H. Kim and Anderson, 2019; Le Pelley et al., 2015; Milstein & Dorris, 2007; Milstein & Dorris, 2011; Theeuwes & Belopolsky, 2012), but the reward association is never assigned to the eye movement itself. In the present study, through a combination of reinforcement-guided selection history and reward feedback, participants were encouraged to repeatedly make saccades towards a target in a particular direction. The training contingencies had a robust influence on eye movements, with 70% of all choice saccades occurring in the reward-associated direction. This saccade preference was driven by a bias to orient in a particular direction rather than to stimuli appearing within a particular region of space, as participants were equally likely to look in the high-value direction regardless of the direction of the initial saccade. This suggests that the high-value direction had greater attentional priority than low-value directions. Because participants were not explicitly informed of the reward structure of the task, this attentional bias was the result of learning from experience.

Despite the irrelevance of reward associations in the test phase, participants were still biased to make initial saccades in the direction previously associated with high reward, but the bias to saccade in the previously-rewarded direction did not persist past the first saccade. That the bias was restricted to the first saccade is not surprising given that there are more potential targets remaining in the directions other than that of the prior choice given the structure of the stimulus displays used (and possibly no stimuli further in the initially-saccaded direction that could be fixated next on some trials, which may explain the tendency to saccade in a direction different from the previously-rewarded direction following the initial saccade).

We employed an unguided visual search task in the test phase to assess the influence of prior learning without competing influences of goal-directed attentional priority, analogous to the free-viewing task of Anderson and Kim (2018a, 2018b). No one element in the display was distinct from the others, providing no basis for feature-based guidance. The resulting search patterns provide direct support for an influence of reward history on the oculomotor system when searching through a display.

## 3. Experiment 2

Experiment 1 demonstrated a bias to direct an initial saccade in the direction previously associated with high reward during unguided visual search (foraging). In Experiment 2, we sought to determine whether a similar bias would be evident in the context of goal-directed visual search. To this end, following the same training procedure,

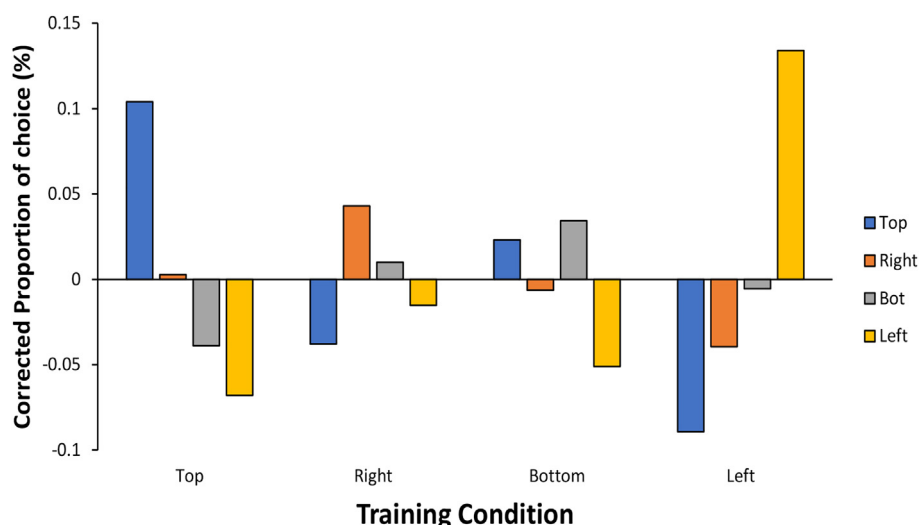
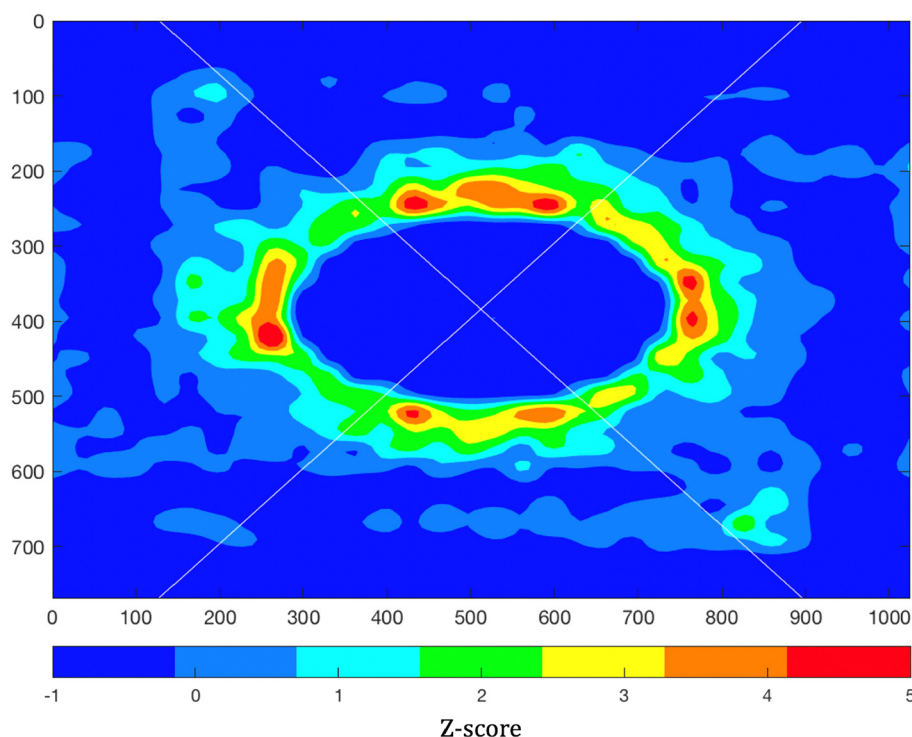


Fig. 2. The distribution of proportion of choices per training condition, for each cardinal direction. The y-axis is the proportion of choice corrected for training-independent bias (see Methods).



**Fig. 3.** A visualization of saccades towards the high-value direction, with direction rotated such that the high-value direction is always to the right for each condition. The X and Y axes reflect pixels on the monitor, and z-scores are computed over the number of fixations to the corresponding pixel (with the resulting z-score map smoothed for visualization).

participants completed a test phase in which the task was to saccade to a shape-defined target. Of interest was whether eye movements would be facilitated on trials in which the target appeared in the previously reward-associated direction, consistent with a bias to execute a saccade in that direction. It cannot be assumed that the oculomotor bias evident in the test phase of Experiment 1 would be sufficiently strong to exert an influence over-and-above goal-directed influences on oculomotor priority in a goal-directed search task, as attentional biases for regions of space do not translate from unguided (free viewing) to guided search when the scenes do not contain objects that could themselves guide attention (Anderson & Kim, 2018a, 2018b).

### 3.1. Methods

#### 3.1.1. Participants

Forty new participants (18–35 years of age,  $M = 22.5$  years; 27 females, 12 males, 1 no response) were recruited from the Texas A&M Community. Participants were compensated with money earned in the experimental task (mean = \$16.15, range = \$10.80–\$20.16). All reported normal or corrected-to-normal visual acuity and normal color vision. All procedures were approved by the Texas A&M University Institutional Review Board and conformed with the principles outlined in the Declaration of Helsinki.

#### 3.1.2. Apparatus

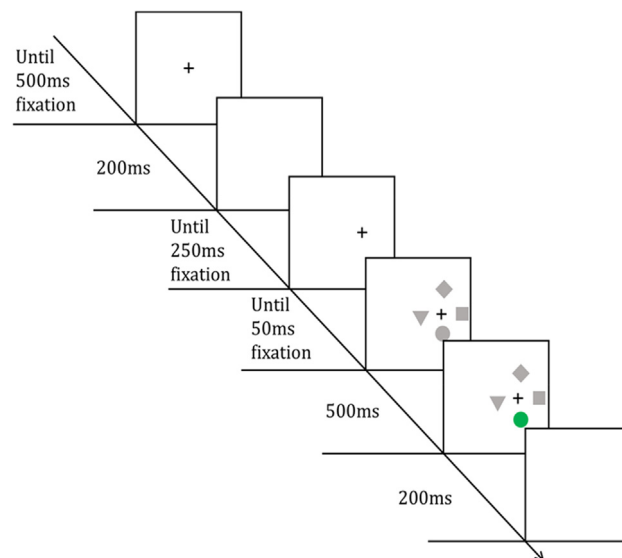
The apparatus was identical to that used in Experiment 1.

#### 3.1.3. Training phase

The training phase was identical to that used in Experiment 1, with the exception that a cross was used instead of a circle for the first-step saccade before the directional choice display, to more closely match the appearance of the stimulus displays used in the test phase.

#### 3.1.4. Test phase

Similar to the training phase, each trial consisted of a fixation display, a secondary fixation display, and a search array (see Fig. 4). The fixation display (cross 1.3°) remained on the screen until 500 ms of continuous fixation within 1.8° of the fixation cross had been registered.



**Fig. 4.** Time course of trial events during the test phase of Experiment 2. Participants fixated on the cross before finding the circular target and fixating on it. The chosen shape would turn red if incorrect, green if correct, or the words “Too Slow!” would appear if no shape was fixated before the timeout limit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

After a 200 ms blank screen, the same cross (with the same fixation window) would appear in one of four cardinal directions at 5.8° eccentricity and remain until eye position had been registered within the fixation window for a continuous period of 250 ms. A search array consisting of four grey shapes jittered (randomly up to 20° of arc in either direction from the cardinal positions) on an imaginary circle with a radius of 7.9° was then presented for 1000 ms or until fixation on a target was registered for a minimum of 150 ms. The nontarget shapes were always a square (2.3°), a diamond (2.2°), and a triangle (3.0°), and the target shape was always a circle (2.85°). A region extending 3.9°



beyond the center of each shape was used as an area of interest for determining stimulus fixations.

The shapes appeared randomly in each cardinal direction, with the constraint that the target circle appeared in each cardinal position equally-often in each block. Participants were instructed to “look at the circle as fast as possible.” If the target was the first shape fixated, it would turn green; otherwise, the non-target shape the participant looked at first would turn red. Correct and incorrect feedback (color change) were presented for 500 ms, while timeout feedback with the words “Too Slow!” at the center of the screen was presented for 1000 ms. Each block consisted of 80 trials, the order of which was randomized.

### 3.1.5. Procedure

The procedure was identical to that of Experiment 1.

### 3.1.6. Analysis of fixations and response times

For the test phase, a trial was considered valid if the target appeared in the previously high-value direction, and invalid if it appeared in any of the other three directions. Only correct responses were included in the mean RT for each participant, and RTs exceeding 2.5 standard deviations (SDs) of the mean for each condition for each participant were trimmed. The RT trimming procedure resulted in the exclusion of 2.6% of trials. The validity effect was calculated by subtracting valid target RTs from invalid target RTs, which were normalized in the same manner as Experiment 1 to account for training-independent biases to more rapidly saccade to a target in a particular direction.

## 3.2. Results and discussion

### 3.2.1. Training phase

Participants were able to learn the reward association, with 64.1% of all saccades going in the high-value direction (as opposed to 25% if participants made saccades in random directions),  $t(39) = 11.565$ ,  $p < 0.001$ ,  $d = 1.829$ . Saccades made in the high-value direction increased over the course of the task, from 43.1% in the first block to 76.1% in the last block,  $t(39) = 8.385$ ,  $p < 0.001$ ,  $d = 1.326$ . Learning rates were overall similar to Experiment 1, as the percent of saccades in the high-value direction did not differ between experiments either in total or in the last block of trials ( $ps > 0.21$ ). To determine whether the learning was robust to the direction of the initial (first-step) saccade, as would be predicted from a directional bias rather than a bias to orient to a particular region of space, we ran a paired samples  $t$ -test comparing the percentage of high-value target choice between first-step saccades in the rewarded direction and in the opposite direction and found no difference between them in any block ( $ps > 0.131$ ). Our results from the training phase closely replicate the findings from Experiment 1.

### 3.2.2. Test phase

Rewarding directional saccades during training produced a 9 ms increase in the validity effect, which was significant,  $t(39) = 2.47$ ,  $p = 0.018$ ,  $d = 0.389$  (see Fig. 5). No corresponding bias was evident in accuracy,  $t(39) = 0.041$ ,  $p = 0.967$ , which was overall high (mean accuracy = 99.3%), indicating that the RT data were not contaminated by a speed-accuracy tradeoff.

The cuing effects observed in RT differed by training condition,  $F(3,36) = 4.07$ ,  $p = 0.014$ ,  $\eta^2 = 0.253$ , being particularly pronounced for the participants for whom the bottom location was associated with high-value ( $p = 0.12$  for the same ANOVA with this condition removed). It was also the case that participants were generally slower to saccade to the bottom location collapsed across training conditions: 471 ms vs. 417, 413, and 419 ms (top, right, and left, respectively),  $ts > 8.75$ ,  $ps < 0.001$ ,  $ds > 1.38$ . Although the reason for the differential effectiveness of the different training conditions is unclear, given the generally slower RTs when saccading in the bottom direction

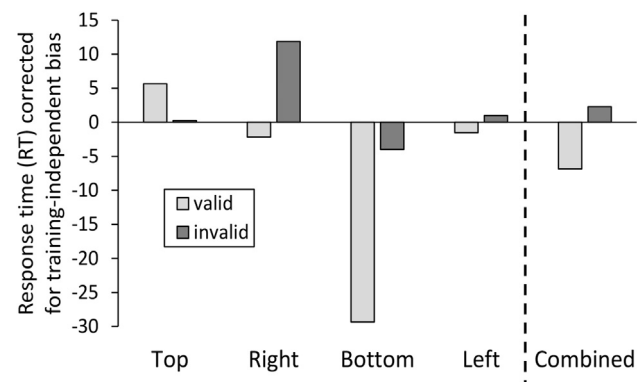


Fig. 5. Response time for valid and invalid trials Experiment 2, corrected for training-independent bias (see Methods), broken down by the different training conditions (high-value direction).

it may be the case that participants simply had more room for improvement in this condition as a result of reward bias. To determine if the observed cuing effects were driven by a general bias towards a particular region in space rather than a directional bias, we ran a  $2 \times 2$  analysis of variance (ANOVA) with first-step saccade (high-value/opposite direction) and validity (valid/invalid) as factors, which revealed no significant interaction between the two,  $F(3,156) = 1.08$ ,  $p = 0.359$ . The results from the test phase of Experiment 2 therefore replicate a bias in directional saccades arising from reward history, in this case extending to a context in which eye movements were also guided by feature-based task goals.

## 4. General discussion

Prior studies have investigated value-driven attentional capture to features such as colors (Anderson et al., 2011; Anderson & Yantis, 2012), shapes (Della Libera et al., 2011; Della Libera & Chelazzi, 2009) and space in both the covert attentional (Anderson, 2015; Chelazzi et al., 2014; Hickey, Chelazzi, & Theeuwes, 2014) and oculomotor (Anderson & Kim, 2018a, 2018b) domains. The findings of the present study extend our understanding of the effects of prior reward learning on attention to the direction of eye movements. We find clear evidence for reward learning modulating saccadic behavior on choices during learning (training phase) and on the execution of directional eye movements during extinction (unrewarded test phase).

In the rewarded training phase, participants made more saccades in the high-value direction and the magnitude of this bias did not differ across first-step saccade direction (i.e., the starting point around which the stimuli were presented on the screen). This suggests that participants learned a directional association instead of associating reward with a particular region of space. In subsequent unguided search without rewards, participants' initial saccade was biased in the direction previous associated with high reward. Likewise, during guided search to a shape-defined target without rewards, participants made faster saccades to targets presented in the previously high-value direction. This RT improvement was also independent of the direction of the first-step saccade, as during training, suggesting that the bias was directional rather than spatial.

The experimental procedure was inspired by that of prior work, specifically Anderson and Kim (2018a, 2018b) where participants formed spatial attentional biases following reward training. Their reward training paradigm was in essence instrumental conditioning, in that reward and choice were linked, and in the present study we applied a similar reward scheme to the direction of saccades. Participants modified their orienting responses in the choice task to better exploit the environment for more reward, indicating robust reward learning. This learning had a broad and persistent influence on the execution of

future eye movements, transferring to both unguided search (a different oculomotor choice task) and guided search (forced-choice orienting). This latter result suggests that the consequence of reward learning on saccades extends beyond the choices participants made and is not specific to a decision-making context.

The magnitude of the learned oculomotor bias might have been different had the rewards during training followed an associative learning procedure within a forced-choice context. If participants had to make an equal number of saccades to each of the four cardinal directions but one direction was more likely to net high reward than low reward, the resulting bias might be comparable to that observed in the present study, driven by the pairing of high reward with the targeting and execution of a directional saccade. Alternatively, this modification could possibly result in poor learning, consistent with work linking reward to the spatial position of targets (Jiang, Sha, & Remington, 2015; Jiang & Swallow, 2013; Jiang, Swallow, Rosenbaum, & Herzig, 2013; Won & Leber, 2016) and with the hypothesis that the manner in which reward learning biases spatial attention is predominantly instrumental and dependent upon selection history (Anderson & Kim, 2018a, 2018b), which contrasts with the influence of reward learning on feature-based attention (Anderson, 2016; H. Kim and Anderson, 2019). Whether the influence of reward learning on directional eye movements is limited to the modulatory role of rewards on selection history or extends to more purely associative reward learning reflects an interesting question to be addressed in future studies.

During training, the position of the targets was randomly jittered around the cardinal directions, resulting in trial-by-trial variation in the specific saccadic trajectory needed to receive high reward. Likewise, during the test phase, the specific position of the stimuli with respect to both the center of the screen and to each other was variable trial-to-trial. That a robust and persistent bias in directional eye movements was still evident under these conditions suggests that the bias is somewhat broadly tuned towards a particular direction. This was particularly evident in the test phase of Experiment 1, given the substantial variability in where stimuli appeared on the screen on each trial and how it differed from the layout of the training phase, attesting to the context generalizability of the bias. Perhaps the positional variability during the training phase facilitated more general learning. The foraging context of Experiment 1 is an arguably more ecologically valid environment to assess biases in saccadic behavior; it can be likened to foraging a newly-encountered bush for berries and having the location of the first berry to be inspected be a function of prior reward learning. It remains to be seen if changing the search context past a threshold could induce exploration (rather than a bias to exploit prior knowledge about rewards; Hills et al., 2015; Wolfe, 2013; Wolfe, Cain, & Alaoui-Soce, 2018; Zhang, Gong, Fougner, & Wolfe, 2017) and bias behavior away from the trained direction, consistent with the role of novelty in guiding attention (Johnston, Hawley, Plewe, Elliott, & DeWitt, 1990; Johnston & Schwarting, 1997).

How this bias is represented in the brain remains an open question. It could be reflected in a spatial representation that is retinotopic and not tied to the layout of a scene, which are known to be involved in directing eye movements (priority maps in parietal cortex, frontal eye field; Bisley & Goldberg, 2010; Desimone & Duncan, 1995; Serences et al., 2005; Sprague & Serences, 2013) and in the execution of oculomotor plans (e.g., superior colliculus; Sommer & Wurtz, 2004). It could also share representational overlap with feature-based value-driven attention and recruit the caudate tail, which is associated with object-selective responses (Beckstead, Edwards, Frankfurter, & Hikosaka, 1981; Griggs et al., 2017; Kim & Hikosaka, 2013; Yamamoto et al., 2013), directing eye movements (Anderson, 2019; Beckstead et al., 1981; Griggs et al., 2017; Kim & Hikosaka, 2013; Yamamoto, Monosov, Yasuda, & Hikosaka, 2012; Yamamoto et al., 2013), and value-dependent modulation (Beckstead et al., 1981; Griggs et al., 2017; Anderson et al., 2016; Anderson, Laurent, & Yantis, 2014; A.J. Kim and Anderson, 2019; Kim & Hikosaka, 2013; Yamamoto et al.,

2013). This would imply that the combination of a stimulus – grey circle – and a particular location evokes a stronger response in this region. The nature of the representation could also be due to semantic (in this case, directional) information being prioritized over the course of learning (Barnard, Scott, Taylor, May, & Knightley, 2004; Fischer, Castel, Dodd, & Pratt, 2003; Krasich, Biggs, & Brockmole, 2018; Krebs, Boehler, Egner, & Woldorff, 2011; Krebs, Boehler, & Woldorff, 2010).

Regardless of the origins of value-driven saccadic biases, the findings of our current study conflict with prior studies that do not report a lasting bias following reward training in an oculomotor task (Dunne et al., 2019; Dunne, Ellison, & Smith, 2015). For example, Dunne et al.'s (2019) study tested the effect of rewarded spatial locations on the oculomotor system. Their paradigm trained participants to make single-step saccades towards differently rewarded targets (with reward varying with the location of the target) but found no facilitation of RT in unrewarded remote distractor and antisaccade tasks. It is possible that reward-training does not generalize to conflict tasks if the new context is too different for any biases to become evident. Another potentially important difference between studies is the choice element of our training phase, which may produce more robust directional biases, consistent with prior work investigating spatial attention (Anderson & Kim, 2018b, 2018a; Jiang et al., 2015, 2013; Jiang & Swallow, 2013; Won & Leber, 2016). It also seems that they were tuning the attentional system to modulate the oculomotor system, which is conceptually different from our paradigm which trains oculomotor selection per se. Our study provides clear evidence that, at least under certain circumstances, it is possible to shape oculomotor behavior with rewards.

## 5. Conclusions

Our findings show that not only can reward influence attention for stimulus features and spatial locations, but also the execution of oculomotor behavior. Biasing patterns of directional saccades could be useful for training more efficient search and information gathering patterns, for example in searching through radiological images to detect cancer (Aizenman, Drew, Ehinger, Georgian-Smith, & Wolfe, 2017; Brennan et al., 2018), performing accurate and efficient baggage screening in security contexts (Kramer, Porfido, & Mitroff, 2019; Mitroff, Ericson, & Sharpe, 2018), or extracting the meaning of a complex graph or data display (Yuan, Haroz, & Franconeri, 2019). With the advent of dynamic visualization, it may become even more advantageous for the directions of saccades to be biased over-and-above feature- and location-based orienting (Xiong, van Weelden, & Franconeri, 2019). The value-driven saccadic bias that we show suggests a novel mechanism by which reward learning influences the visual system, which could be supported by distinct learning systems in the brain.

## Funding statement

The reported research was supported in part by a start-up package to B.A.A. from Texas A&M University and NARSAD Young Investigator Grant 26008 to B.A.A. and NIH grant R01-DA046410 to B.A.A.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2019.104145>.

## References

- Aizenman, A., Drew, T., Ehinger, K. A., Georgian-Smith, D., & Wolfe, J. M. (2017). Comparing search patterns in digital breast tomosynthesis and full-field digital mammography: An eye tracking study. *Journal of Medical Imaging (Bellingham, Wash.)*, 4(4), 045501. <https://doi.org/10.1117/1.JMI.4.4.045501>
- Anderson, B. A. (2015). Value-driven attentional capture is modulated by spatial context. *Visual Cognition*, 23(1–2), 67–81. <https://doi.org/10.1080/13506285.2014.956851>.

- Anderson, B. A. (2016). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences*, 1369(1), 24–39. doi:<https://doi.org/10.1111/nyas.12957>
- Anderson, B. A. (2019). Neurobiology of value-driven attention. *Current Opinion in Psychology*. <https://doi.org/10.1016/j.copsyc.2018.11.004>.
- Anderson, B. A., & Kim, H. (2018a). Mechanisms of value-learning in the guidance of spatial attention. *Cognition*, 178, 26–36. <https://doi.org/10.1016/j.COgnITION.2018.05.005>.
- Anderson, B. A., & Kim, H. (2018b). On the representational nature of value-driven spatial attentional biases. *Journal of Neurophysiology*. <https://doi.org/10.1152/jn.00489.2018>.
- Anderson, B. A., Kuwabara, H., Wong, D. F., Gean, E. G., Rahmim, A., Brašić, J. R., ... Yantis, S. (2016). The role of dopamine in value-based attentional orienting. *Current Biology*. <https://doi.org/10.1016/j.cub.2015.12.062>.
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, 108(25), 10367–10371. <https://doi.org/10.1073/PNAS.1104047108>.
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2014). Value-driven attentional priority signals in human basal ganglia and visual cortex. *Brain Research*, 1587, 88–96. <https://doi.org/10.1016/j.brainres.2014.08.062>.
- Anderson, B. A., & Yantis, S. (2012). Value-driven attentional and oculomotor capture during goal-directed, unconstrained viewing. *Attention, Perception, and Psychophysics*. <https://doi.org/10.3758/s13414-012-0348-2>.
- Barnard, P. J., Scott, S., Taylor, J., May, J., & Knightley, W. (2004). Paying attention to meaning. *Psychological Science*, 15(3), 179–186. <https://doi.org/10.1111/j.0956-7976.2004.01503006.x>.
- Beckstead, R. M., Edwards, S. B., Frankfurter, A., & Hikosaka, O. (1981). A comparison of the intranigral distribution of nigroreticular neurons labeled with horseradish peroxidase in the monkey, cat, and rat. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 1(2), 121–125. <https://doi.org/10.1523/jneurosci.0828-12.2012>.
- Bendiksy, M. S., & Platt, M. L. (2006). Neural correlates of reward and attention in macaque area LIP. *Neuropsychologia*, 44(12), 2411–2420. <https://doi.org/10.1016/j.neuropsychologia.2006.04.011>.
- Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*. <https://doi.org/10.1146/annurev-neuro-060909-152823>.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*. <https://doi.org/10.1163/156856897X00357>.
- Brennan, P. C., Gandomkar, Z., Ekpo, E. U., Tapia, K., Trieu, P. D., Lewis, S. J., ... Evans, K. K. (2018). Radiologists can detect the ‘gist’ of breast cancer before any overt signs of cancer appear. *Scientific Reports*, 8(1), 8717. <https://doi.org/10.1038/s41598-018-26100-5>.
- Bucker, B., Silvius, J. D., Donk, M., & Theeuwes, J. (2015). Reward modulates oculomotor competition between differently valued stimuli. *Vision Research*, 108, 103–112. <https://doi.org/10.1016/j.visres.2015.01.020>.
- Chelazzi, L., Eštočinová, J., Calletti, R., Gerfo, E. L., Sani, I., Libera, C. D., ... Santandrea, E. (2014). Altering spatial priority maps via reward-based learning. *Journal of Neuroscience*, 34(25), 8594–8604. <https://doi.org/10.1523/JNEUROSCI.0277-14.2014>.
- Chen, L. L., Chen, Y. M., Zhou, W., & Mustain, W. D. (2014). Monetary reward speeds up voluntary saccades. *Frontiers in Integrative Neuroscience*, 8, 48. <https://doi.org/10.3389/fnint.2014.00048>.
- Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science*, 20(6), 778–784. <https://doi.org/10.1111/j.1467-9280.2009.02360.x>.
- Della Libera, C., Perlati, A., & Chelazzi, L. (2011). Dissociable effects of reward on attentional learning: From passive associations to active monitoring. *PLoS One*, 6(4), e19460. <https://doi.org/10.1371/journal.pone.0019460>.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18(April), 193–222. <https://doi.org/10.1126/science.1247003>.
- Donohue, S. E., Hopf, J.-M., Bartsch, M. V., Schoenfeld, M. A., Heinze, H.-J., & Woldorff, M. G. (2016). The rapid capture of attention by rewarded objects. *Journal of Cognitive Neuroscience*, 28(4), 529–541. [https://doi.org/10.1162/jocn\\_a.00917](https://doi.org/10.1162/jocn_a.00917).
- Dunne, S., Ellison, A., & Smith, D. T. (2015). Rewards modulate saccade latency but not exogenous spatial attention. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2015.01080>.
- Dunne, S., Ellison, A., Smith, D. T., Dunne, S., Ellison, A., & Smith, D. T. (2019). The limitations of reward effects on saccade latencies: An exploration of task-specificity and strength. *Vision*, 3(2), 20. <https://doi.org/10.3390/vision3020020>.
- Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*. <https://doi.org/10.1038/nn1066>.
- Griggs, W. S., Kim, H., Ghazizadeh, A., Gabriela Costello, M., Wall, K. M., & Hikosaka, O. (2017). Flexible and stable value coding areas in caudate head and tail receive anatomically distinct cortical and subcortical inputs. *Frontiers in Neuroanatomy*, 11, 106. <https://doi.org/10.3389/fnana.2017.00106>.
- Hickey, C., Chelazzi, L., & Theeuwes, J. (2014). Reward-priming of location in visual search. *PLoS One*, 9(7), e103372. <https://doi.org/10.1371/journal.pone.0103372>.
- Hickey, C., Kaiser, D., & Peelen, M. V. (2015). Reward guides attention to object categories in real-world scenes. *Journal of Experimental Psychology: General*, 144(2), 264–273. <https://doi.org/10.1037/a0038627>.
- Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., & Group, the C. S. R. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, 19(1), 46. <https://doi.org/10.1016/j.TICS.2014.10.004>.
- Ikeda, T., & Hikosaka, O. (2003). Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron*. [https://doi.org/10.1016/S0896-6273\(03\)00464-1](https://doi.org/10.1016/S0896-6273(03)00464-1).
- Jiang, Y. V., Sha, L. Z., & Remington, R. W. (2015). Modulation of spatial attention by goals, statistical learning, and monetary reward. *Attention, Perception, & Psychophysics*, 77(7), 2189–2206. <https://doi.org/10.3758/s13414-015-0952-z>.
- Jiang, Y. V., & Swallow, K. M. (2013). Spatial reference frame of incidentally learned attention. *Cognition*, 126(3), 378–390. <https://doi.org/10.1016/J.COgnITION.2012.10.011>.
- 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. <https://doi.org/10.1037/a0027611>.
- Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, J. M. G., & DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/0096-3445.119.4.397>.
- Johnston, W. A., & Schwartz, I. S. (1997). Novel popout: An enigma for conventional theories of attention. *Journal of Experimental Psychology: Human Perception and Performance*. <https://doi.org/10.1037/0096-1523.23.3.622>.
- Kawagoe, R., Takikawa, Y., & Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neuroscience*, 1(5), 411–416. <https://doi.org/10.1038/1625>.
- Kim, A. J., & Anderson, B. A. (2019a). Neural correlates of attentional capture by stimuli previously associated with social reward. *Cognitive Neuroscience*, 1–11. doi:<https://doi.org/10.1080/17588928.2019.1585338>.
- Kim, H., & Anderson, B. A. (2019b). Dissociable components of experience-driven attention. *Current Biology*, 29(5), 841–845.e2. doi:<https://doi.org/10.1016/j.cub.2019.01.030>.
- Kim, H. F., & Hikosaka, O. (2013). Distinct basal ganglia circuits controlling behaviors guided by flexible and stable values. *Neuron*, 79(5), 1001–1010. <https://doi.org/10.1016/J.NEURON.2013.06.044>.
- Kramer, M. R., Porfido, C. L., & Mitroff, S. R. (2019). Evaluation of strategies to train visual search performance in professional populations. *Current Opinion in Psychology*. <https://doi.org/10.1016/j.copsyc.2019.01.001>.
- Krasich, K., Biggs, A. T., & Brockmole, J. R. (2018). Attention capture during visual search: The consequences of distractor appeal, familiarity, and frequency. *Visual Cognition*, 1–19. <https://doi.org/10.1080/13506285.2018.1508102>.
- Krebs, R. M., Boehler, C. N., Egner, T., & Woldorff, M. G. (2011). The neural underpinnings of how reward associations can both guide and misguide attention. *Journal of Neuroscience*. <https://doi.org/10.1523/jneurosci.0732-11.2011>.
- Krebs, R. M., Boehler, C. N., & Woldorff, M. G. (2010). The influence of reward associations on conflict processing in the Stroop task. *Cognition*. <https://doi.org/10.1016/j.cognition.2010.08.018>.
- Laurent, P. A., Hall, M. G., Anderson, B. A., & Yantis, S. (2015). Valuable orientations capture attention. *Visual Cognition*, 23(1–2), 133–146. <https://doi.org/10.1080/13506285.2014.965242>.
- Le Pelley, M. E., Pearson, D., Griffiths, O., & Beesley, T. (2015). When goals conflict with values: Counterproductive attentional and oculomotor capture by reward-related stimuli. *Journal of Experimental Psychology: General*, 144(1), 158–171. <https://doi.org/10.1037/xge0000037>.
- Milstein, D. M., & Dorris, M. C. (2007). The influence of expected value on saccadic preparation. *Journal of Neuroscience*, 27(18), 4810–4818. <https://doi.org/10.1523/JNEUROSCI.0577-07.2007>.
- Milstein, D. M., & Dorris, M. C. (2011). The relationship between saccadic choice and reaction times with manipulations of target value. *Frontiers in Neuroscience*, 5, 122. <https://doi.org/10.3389/fnins.2011.00122>.
- Mitroff, S. R., Ericson, J. M., & Sharpe, B. (2018). Predicting airport screening officers' visual search competency with a rapid assessment. *Human Factors*. <https://doi.org/10.1177/0018720817743886>.
- NisSENS, T., Failing, M., & Theeuwes, J. (2016). People look at the object they fear: oculomotor capture by stimuli that signal threat. *Cognit. Emot.* 31(8), 1707–1714. <https://doi.org/10.1080/02699931.2016.1248905>.
- Purves, D., Augustine, G. J., & Fitzpatrick, D. (2012). *Neuroscience* (5th Edition). *Nature reviews neuroscience* <https://doi.org/978-0878937257>.
- Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science*, 16(2), 114–122. <https://doi.org/10.1111/j.0956-7976.2005.00791.x>.
- Sohn, J., & Lee, D. (2006). Effects of reward expectancy on sequential eye movements in monkeys. *Neural Networks*, 19(8), 1181–1191. <https://doi.org/10.1016/J.NEUNET.2006.04.005>.
- Sommer, M. A., & Wurtz, R. H. (2004). What the brain stem tells the frontal cortex. I. Oculomotor signals sent from superior colliculus to frontal eye field via mediodorsal thalamus. *Journal of Neurophysiology*. <https://doi.org/10.1152/jn.00738.2003>.
- Sprague, T. C., & Serences, J. T. (2013). Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nature Neuroscience*, 16(12), 1879–1887. <https://doi.org/10.1038/nn.3574>.
- Takikawa, Y., Kawagoe, R., & Hikosaka, O. (2002). Reward-dependent spatial selectivity of anticipatory activity in monkey caudate neurons. *Journal of Neurophysiology*, 87(1), 508–515. <https://doi.org/10.1152/jn.00288.2001>.
- Theeuwes, J., & Belopolsky, A. V. (2012). Reward grabs the eye: Oculomotor capture by rewarding stimuli. *Vision Research*, 74, 80–85. <https://doi.org/10.1016/j.visres.2012.07.024>.
- Wolfe, J. M. (2013). When is it time to move to the next raspberry bush? Foraging rules in human visual search. *Journal of Vision*, 13(3), 10. <https://doi.org/10.1167/13.3.10>.
- Wolfe, J. M., Cain, M. S., & Alaoui-Socé, A. (2018). Hybrid value foraging: How the value of targets shapes human foraging behavior. *Attention, Perception, & Psychophysics*,



- 80(3), 609–621. <https://doi.org/10.3758/s13414-017-1471-x>.
- Won, B.-Y., & Leber, A. B. (2016). How do magnitude and frequency of monetary reward guide visual search? *Attention, Perception, & Psychophysics*, 78(5), 1221–1231. <https://doi.org/10.3758/s13414-016-1154-z>.
- Xiong, C., van Weelden, L., & Franconeri, S. (2019). The curse of knowledge in visual data communication. *IEEE Transactions on Visualization and Computer Graphics*. <https://doi.org/10.1109/tvcg.2019.2917689>.
- Yamamoto, S., Kim, H. F., & Hikosaka, O. (2013). Reward value-contingent changes of visual responses in the primate caudate tail associated with a visuomotor skill. *Journal of Neuroscience*, 33(27), 11227–11238. <https://doi.org/10.1523/JNEUROSCI.0318-13.2013>.
- Yamamoto, S., Monosov, I. E., Yasuda, M., & Hikosaka, O. (2012). What and where information in the caudate tail guides saccades to visual objects. *Journal of Neuroscience*. <https://doi.org/10.1523/jneurosci.0828-12.2012>.
- Yuan, L., Haroz, S., & Franconeri, S. (2019). Perceptual proxies for extracting averages in data visualizations. *Psychonomic Bulletin and Review*. <https://doi.org/10.3758/s13423-018-1525-7>.
- Zhang, J., Gong, X., Fougner, D., & Wolfe, J. M. (2017). How humans react to changing rewards during visual foraging. *Attention, Perception, & Psychophysics*, 79(8), 2299–2309. <https://doi.org/10.3758/s13414-017-1411-9>.