Cognition 178 (2018) 26–36

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Mechanisms of value-learning in the guidance of spatial attention

Brian A. Anderson*, Haena Kim

Texas A&M University, United States

ARTICLE INFO

Keywords:
Selective attention
Spatial attention
Attentional capture
Reward learning
Real world scenes

ABSTRACT

The role of associative reward learning in the guidance of feature-based attention is well established. The extent to which reward learning can modulate spatial attention has been much more controversial. At least one demonstration of a persistent spatial attention bias following space-based associative reward learning has been reported. At the same time, multiple other experiments have been published failing to demonstrate enduring attentional biases towards locations at which a target, if found, yields high reward. This is in spite of evidence that participants use reward structures to inform their decisions where to search, leading some to suggest that, unlike feature-based attention, spatial attention may be impervious to the influence of learning from reward structures. Here, we demonstrate a robust bias towards regions of a scene that participants were previously rewarded for selecting. This spatial bias relies on representations that are anchored to the configuration of objects within a scene. The observed bias appears to be driven specifically by reinforcement learning, and can be observed with equal strength following non-reward corrective feedback. The time course of the bias is consistent with a transient shift of attention, rather than a strategic search pattern, and is evident in eye movement patterns during free viewing. Taken together, our findings reconcile previously conflicting reports and offer an integrative account of how learning from feedback shapes the spatial attention system.

1. Introduction

The role of an observer's goals (*top-down* factors) and the physical salience of objects (*bottom-up* factors) in the control of attention have been well established and serve as the foundation for prominent models of selective attention (e.g., Corbetta & Shulman, 2002; Desimone & Duncan, 1995; Theeuwes, 2010; Wolfe, Cave, & Franzel, 1989). More recently, it has been argued that this dichotomy cannot explain the role of selection history in the control of attention, which appears to be both non-strategic and independent of the physical salience of stimuli (Awh, Belopolsky, & Theeuwes, 2012). In this context, an important component of selection history has been argued to reflect associative reward learning, with objects previously associated with reward automatically capturing visual attention (Anderson, 2013).

The role of associative reward learning in the control of attention was initially demonstrated using stimuli defined by shape, with results showing that stimulus competition was biased for or against different shapes based on whether observers were rewarded for selecting or ignoring them, respectively (Della Libera & Chelazzi, 2009). This bias carried over into extinction, suggesting that it was non-strategic. Powerful evidence for the unique role of associative reward learning in the control of attention was provided by a study in which task-irrelevant distractors were rendered in a color that had been predictive of reward during a prior training phase. These distractors were not physically salient (less so than the target), and the color of stimuli was known by participants to be completely irrelevant to the task. Nevertheless, attention was automatically captured by the previously reward-associated colors, suggesting a distinct mechanism of attentional control that has been referred to as *value-driven attention* (Anderson, Laurent, & Yantis, 2011).

Many subsequent studies have adopted this approach of associating stimulus features (often color) with reward and then presenting the previously reward-associated features as distractors, replicating and extending the phenomenon of value-driven attention (e.g., Anderson, 2016a, 2016c; Anderson, Folk, Garrison, & Rogers, 2016; Anderson, Laurent, & Yantis, 2012; Anderson & Yantis, 2012, 2013; Anderson, Kuwabara, et al., 2016; Failing & Theeuwes, 2014; Le Pelley, Pearson, Griffiths, & Beesley, 2015; Mine & Saiki, 2015; Moher, Anderson, & Song, 2015; Pool, Brosch, Delplanque, & Sander, 2014; see Anderson, 2016b, for a recent review). Attention has been successfully trained to favor a variety of stimulus features, ranging from specific colors (e.g., Anderson et al., 2011) and orientations (Laurent, Hall, Anderson, & Yantis, 2015; Lee & Shomstein, 2014) to shapes (Della Libera & Chelazzi, 2009) and even object categories (Hickey & Peelen, 2015). In addition to feature-based attention, object-based attention also appears to be strongly modulated by associative reward learning (Lee &

* Corresponding author at: Texas A&M University, Department of Psychology, 4235 TAMU, College Station, TX 77843-4235, United States.
E-mail address: brian.anderson@tamu.edu (B.A. Anderson).

https://doi.org/10.1016/j.cognition.2018.05.005
Received 13 October 2017; Received in revised form 24 April 2018; Accepted 5 May 2018

Shomstein, 2013; Shomstein & Johnson, 2013).

Few studies have probed the influence of associative reward learning on the control of space-based attention, where reward is not predicted by a particular stimulus feature but rather by where in space attention needs to be directed in order to receive high reward. Value-driven attentional biases for a stimulus feature can be modulated by spatial information, such that the bias is specific to when that feature appears in a region of space in which it was rewarded (Anderson, 2015a). More purely space-based attentional biases have also been shown to be modulated by reward. When a high reward is received for identifying a target appearing in a given location, this location is prioritized on the subsequent trial (Hickey, Chelazzi, & Theeuwes, 2014), extending earlier evidence for reward-mediated priming of stimulus color (Hickey, Chelazzi, & Theeuwes, 2010). A more enduring bias towards a previously rewarded location was demonstrated following a multi-day training protocol in which participants performed a difficult visual search for alphanumeric among nonalphanumeric characters. Which of eight possible stimulus positions a searched-for character appeared in predicted the amount of money earned for reporting that target on a given trial. During extinction, participants were more likely to report a target appearing in a previously high-value location, specifically when two targets were simultaneously presented that competed for attention (Chelazzi et al., 2014).

At least two cases have been reported in which a spatial reward manipulation failed to produce any evidence for an enduring attentional bias, or even an attentional bias during the period in which the reward structure was currently in place. In each study, multiple experiments were conducted in which participants searched for a "T" among offset "L" distractors (Jiang, Li, & Remington, 2015; Won & Leber, 2016). When the target appeared anywhere within a particular quadrant of the screen, it was much more likely to yield a high reward if correctly identified. Under a variety of conditions, including conditions of time pressure in which participants should be highly motivated to preferentially search the high-value quadrant in order to maximize rewards, no measurable spatial attention bias was observed (Jiang et al., 2015; Won & Leber, 2016). Furthermore, under similar conditions in which participants were asked to instead choose a particular stimulus rather than perform visual search, robust spatial preferences for highly rewarded locations were observed. These results suggest that spatial reward can readily influence choice behavior (Won & Leber, 2016), but seemingly not the allocation of attention during visual search (Jiang et al., 2015; Won & Leber, 2016).

It is important to note that space-based attentional biases are robustly influenced by a different source of selection history using the same experimental paradigm that failed to show value-driven biases. Specifically, participants are much faster to report targets in locations that more frequently contained targets in prior trials, even well after such biased probabilities are no longer in place (e.g., Jiang & Swallow, 2013; Jiang, Swallow, Rosenbaum, & Herzig, 2013; Jiang et al., 2015; Won & Leber, 2016). Similarly, targets can be found more quickly when the position of the target is consistently predicted by the spatial configuration of non-targets (e.g., Chun & Jiang, 1998, 2003), a phenomenon termed *contextual cueing*. That a different form of selection history can so robustly bias spatial attention in this paradigm argues against a general insensitivity of the paradigm to the ability to detect a learned spatial attention bias.

These repeated failures to observe reliable effects of reward history on the allocation of spatial attention during visual search have naturally led to skepticism concerning whether principles of value-driven attention extend to the spatial domain (Jiang et al., 2015; Won & Leber, 2016). Indeed, it has been suggested that evolutionary pressures imposed by naturally occurring reward structures might strongly favor feature-reward pairings over space-reward pairings, rendering influences of the reward system on spatial attention phylogenetically implausible (Won & Leber, 2016). We would argue that this is a fair point, in the context of how space is defined in these studies.

In traditional visual search paradigms, including those used by Jiang et al. (2015) and Won and Leber (2016), space is defined in a highly abstract manner: a region of a blank computer screen. In fact, there is no clear anchor point for defining where one region would end and another begin, apart from the borders imposed by the edges of the monitor. Such a highly abstract notion of space is unlikely to engage the spatial representations one might use to guide search for a valued item based on learning history, such as where ice cream tends to be stored in the freezer. In this case, the valued location is defined in the context of the spatial arrangement of objects in the scene (e.g., the position of the freezer relative to other objects in the room, and which section of which shelf when looking inside the freezer). The spatial information provided by real-world scenes can serve as the basis for contextual cueing of target position (e.g., Brockmole & Henderson, 2006a, 2006b), suggesting a rich source of spatial guidance, although contextual cueing is also evident with the more abstract stimulus displays that have failed to produce evidence of value-based attentional biases (Jiang et al., 2015; Won & Leber, 2016).

Perhaps information pertaining to the spatial layout and arrangement of objects in a scene is useful for guiding spatial attention on the basis of reward history, which might help explain the apparent discrepancy between Chelazzi et al. (2014) on the one hand, and Jiang et al. (2015) and Won and Leber (2016) on the other hand. Chelazzi et al. (2014) only found an effect of reward when two targets simultaneously competed for attention, where reward was not only tied to the absolute spatial location of the targets but also to their relative positions. With the aim of reconciling these conflicting reports, in the present study, we examined the role of value learning in the context of real-world scenes, both scenes containing a rich array of objects with a consistent spatial arrangement and scenes containing no objects (textures).

2. Experiment 1

In Experiment 1, participants were first trained to associate a specific region of multiple different scenes with monetary reward. On each trial, one of eight scenes remained on the screen until participants clicked on a pixel within the scene using the mouse cursor. Participants were instructed that they would be rewarded for each click, and that the amount of reward received depended on where they clicked. Unbeknownst to the participant, for each of the eight scenes, clicking in one quadrant would always yield more reward than clicking in any other quadrant, and clicking in the center of that quadrant was associated with the best possible payout. Each quadrant served as the high-value quadrant equally-often across scenes, requiring that participants' memory for high-value locations be context-specific, rather than reflect a global bias towards one particular region of the computer screen. In the feature domain, value-driven attentional capture can exhibit contextual specificity (Anderson, 2015b). In the present study, the scene context manipulation demanded that participants take into account the unique spatial layout of each scene.

In a subsequent test phase, participants performed visual search for a side-ways "T" among three upright or upside down "T" distractors, with one search item appearing in the center of each quadrant on the screen. The previously presented scenes were used as the background and were irrelevant to the task, and participants were informed that they could neither earn nor lose money in this task. To the degree that spatial attention is automatically oriented towards previously high-value locations within a scene, participants should be significantly faster to report the target when it appears within a previously high-value location, which would be reflected in a robust validity effect.

2.1. Methods

2.1.1. Participants

Thirty-six participants were recruited from the Texas A&M

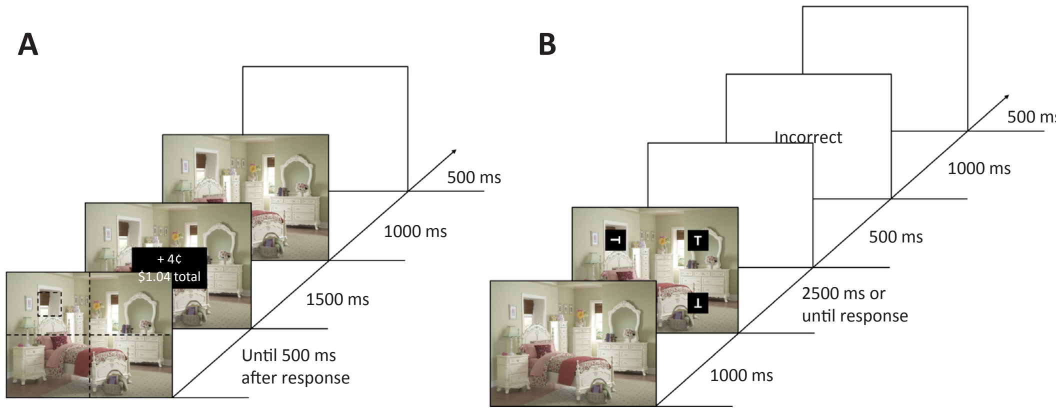

Fig. 1. Time course of trial events during the training phase and test phase of Experiment 1. (A) Training phase. Participants clicked on a pixel each time a scene was presented, and were rewarded for the click contingent upon which quadrant of the picture they had clicked on. In this example, the upper-left quadrant is the high-value quadrant, and clicking within the center box in that quadrant yielded the highest possible reward. The dotted lines were not visible to participants and are provided for illustrative purposes. (B) Test phase. Participants searched for a side-ways "T" among upright and upside down "T" distractors. Scenes previously experienced during the training phase were used as the background and were irrelevant to the task. Note that the stimuli are not drawn to scale in the figure.

University community. Participants were compensated with money earned in the experimental task. All reported normal or corrected-to-normal visual acuity and normal color vision. Data from one participant was replaced due to chance-level performance in the test phase (accuracy = 51%). All procedures were approved by the Texas A&M University Institutional Review Board and conformed with the principles outlined in the Declaration of Helsinki.

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. Manual responses were entered using a standard keyboard.

2.1.3. Training phase

Each trial began with the presentation of a scene image that filled the entire computer screen (see Fig. 1A). The scene image remained on screen until participants clicked on the scene using the mouse cursor. 500 ms after a click was registered, feedback was presented at the center of the screen within a black box for 1500 ms. The feedback indicated the money earned for the preceding click, along with the current bank total. The feedback then disappeared while the scene remained on screen for an additional 1000 ms, which was followed by a blank 500 ms inter-trial-interval (ITI).

Participants were informed that they would earn money each time they clicked on a scene, and that how much they earned for each click would depend on where they clicked. Participants were encouraged to maximize their earnings by clicking on good locations, but were not provided any explicit information about where those good locations were.

Eight difference scenes were presented, 40 times each for a total of 320 trials. The scenes used were taken from the CB Database (Sareen, Ehinger, & Wolfe, 2016). For each scene, one quadrant was designated as the high-value quadrant. Any click within the high-value quadrant yielded at least a 4¢ reward. If participants clicked within an imaginary 2.7° × 2.7° box centered within the high-value quadrant, 7¢ was earned. Any click within one of the other three (low-value) quadrants yielded a 1¢ reward.

Participants were assigned to one of four training conditions in counterbalanced fashion, with each quadrant of each scene serving as the high-value quadrant in one of the four conditions. Therefore, across participants, each quadrant served as the high-value quadrant equally-often. The order in which the scenes were presented to each participant was randomized.

2.1.4. Test phase

Each trial began with the presentation of one of the scenes from training for 1000 ms, followed by the presentation of a "T" stimulus within a 2.1° × 2.1° black box centered within each of the four quadrants (see Fig. 1B). One "T" was tilted either 90° to the left or right and served as the target, while the other three "T"s were either upright or upside down (randomly determined with the constraint that all three non-target "T"s could not be oriented in the same direction). The "T" display remained on screen until a response was registered or 2500 ms had elapsed, after which the trial timed out. If participants responded incorrectly, the word "Incorrect" was presented against the center of the screen for 1000 ms following a blank 500 ms interval, whereas if a response was not registered within the 2500 ms response window, the feedback instead read "Too slow." Feedback was omitted following correct responses. Each trial ended with a blank 500 ms ITI.

The test phase consisted of 320 trials, with a 30 sec break occurring after each epoch of 80 trials. Each scene was presented 40 times, with the target appearing in each quadrant of each scene equally-often. The target was tilted 90° to the left and right equally-often for each scene. Trials were presented in a random order.

If the bottom of the target pointed to the right, participants were instructed to press the "m" key, and if the bottom of the target pointed to the left, participants were instructed to press the "z" key. The task began with 10 practice trials that used different scenes than those that were experienced during training. Participants were instructed to respond both fast and accurately, and were informed that money could not be neither gained nor lost in this part of the experiment. Upon completion of the experiment, participants were paid their cumulative earnings from training.

2.1.5. Data analysis

In the training phase, the reward earned on each trial was categorized in terms of how many times the scene on that trial had been presented. This yielded earnings over 40 different scene presentations, with eight different scenes contributing to each value. In the test phase, response time (RT) was recorded from the onset of the four items comprising the search array, and RTs exceeding 3 SD of the mean of their respective condition or faster than 200 ms were trimmed.

2.2. Results and discussion

2.2.1. Training phase

The amount earned differed across scene presentation, $F(39,1365) = 56.82$, $p < 0.001$, $\eta_p^2 = 0.619$ (Fig. 2), with performance beginning at chance and plateauing around the 33rd time each scene was

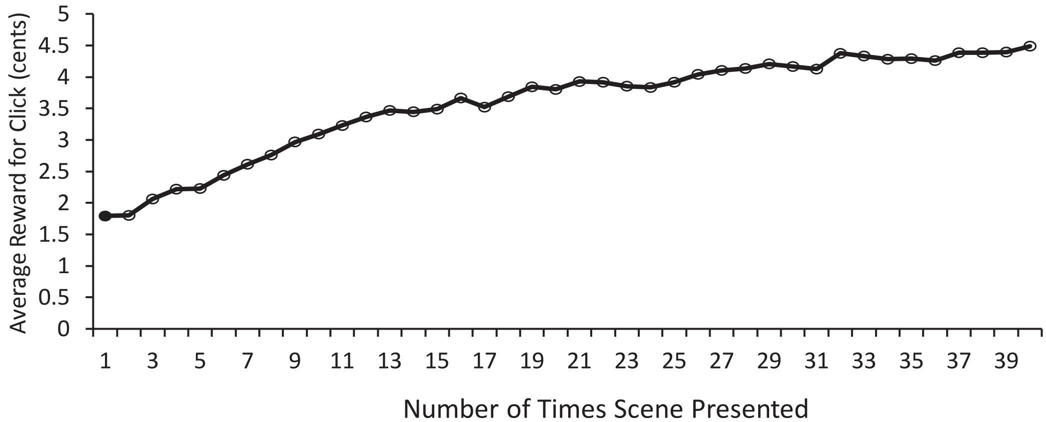

Fig. 2. Reward earned across different presentations of each scene during the training phase of Experiment 1.

presented. The learning curve was well accounted for by a linear trend, $F(1,35) = 147.37$, $p < 0.001$, $\eta_p^2 = 0.808$. Consistent with Won and Leber (2016), our results show that spatial choices were robustly influenced by reward feedback, reflecting more optimal choices with reinforcement learning. Raw data for each phase of all experiments is included as Supplemental Material (File S1).

2.2.2. *Test phase*

A robust cuing effect of 18 ms was evident in the test phase in RT, $t(35) = 2.61$, $p = 0.013$, $d = 0.44$ (Fig. 3). Accuracy was generally high and did not exhibit a cuing effect (valid = 93.4%, invalid = 92.6%), $t(35) = 1.10$, $p = 0.280$, and there was no hint of a speed-accuracy tradeoff as participants were numerically less accurate on invalid trials. When a previously experienced scene served as the background for a visual search task, perception was biased in favor of stimuli appearing within the region of that particular scene that the participant had been more highly rewarded for selecting in the past.

Unlike Jiang et al. (2015) and Won and Leber (2016), robust spatial biases were evident following reward learning when reward was anchored to locations within a scene rather than a location on a blank computer screen. Furthermore, the observed spatial bias was bound specifically to the scene context, as no one viewer-centric location (e.g., upper right) was generally more likely to be rewarded than any other, attesting to the ecological validity of the bias (see Won & Leber, 2016). This contextual specificity may help to explain why a significant spatial bias was observed in Chelazzi et al. (2014) on two-target trials but not single target trials, whereas Jiang et al. (2015) and Won and Leber (2016) failed to find evidence for an influence of reward learning on spatial attention (see *7. General Discussion* for further discussion on this issue). The results provide clear evidence for a role for reward learning in the shaping of spatial attention.

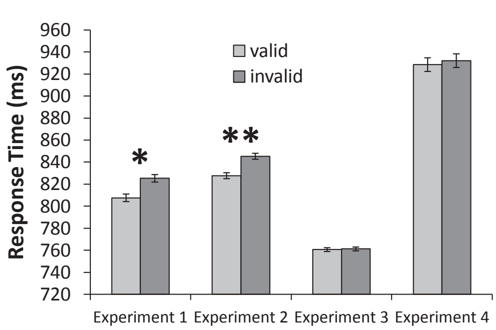

Fig. 3. Response time for valid and invalid trials in Experiments 1–4. Error bars reflect the within-subjects confidence intervals. $^{*}p < 0.05$, $^{**}p < 0.01$.

3. Experiment 2

The results of Experiment 1 provide straightforward evidence that selection history can robustly bias spatial attention in a context-specific manner. However, it is unclear to what degree this observed bias is the product of a reward-mediated process specifically. This is because the amount of reward participants earn on each trial is entirely contingent on their orienting behavior. Not only do participants experience greater rewards when selecting the high-value quadrant, they simply select it more often. This repeated selection, motivated by reinforcement learning, could be more directly related to the observed spatial bias than the rewards that motivate such selection per se.

To the degree that the bias observed in Experiment 1 is the result of selection history rather than reward history per se, it should also be evident following non-reward corrective feedback that similarly motivates participants to repeatedly select a particular quadrant within a scene. To determine whether this is the case, we conducted a second experiment in which, rather than receive different amounts of reward, participants received varying non-reward performance feedback. If participants selected what was a low-value location in the prior experiment, 'Not good' was presented on the computer screen in place of the monetary feedback. If they selected the high-value quadrant off-center, 'Good!' feedback was provided, and if in the center of the high-value quadrant, the word 'Excellent!!' appeared. To the degree that such non-reward feedback is effective in curbing spatial choice behavior, resulting in a robust learning curve as in Experiment 1, a fair test of the necessity of reward in producing an enduring spatial attention bias can be provided.

3.1. *Methods*

3.1.1. *Participants*

Thirty-six new participants were recruited from the Texas A&M University community. Participants were compensated with course credit. All reported normal or corrected-to-normal visual acuity and normal color vision. Data from one participant was replaced due to chance-level performance in the test phase (accuracy = 49%).

3.1.2. *Experimental task*

The experimental task was identical to that used in Experiment 1, with the exception that the feedback consisted of 'Not good', 'Good!', and 'Excellent!!' in place of 1, 4 and 7 cents along with cumulative earnings, respectively. The instructions emphasized using feedback to click on the “best” locations. Instructions for the test phase again emphasized that participants should aim to respond both quickly and accurately, but did not make reference to the absence of monetary rewards as such rewards were not present during training.

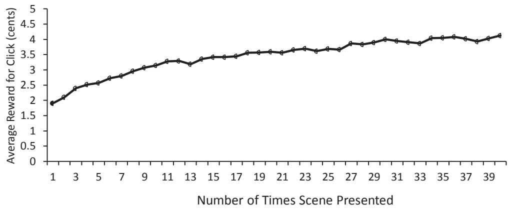

Fig. 4. The quality of feedback received, expressed in numerical terms corresponding to the reward structure used in Experiment 1, across different presentations of each scene during the training phase of Experiment 2.

3.1.3. Data analysis

To facilitate statistical analysis and interpretation of the data across experiments, performance feedback received during the training phase was quantified using the corresponding numerical values from Experiment 1 (i.e., 1, 4, and 7 for 'Not good', 'Good!', and 'Excellent!!', respectively).

3.2. Results and discussion

3.2.1. Training phase

As in Experiment 1, performance differed across scene presentation, $F(39,1365) = 24.06$, $p < 0.001$, $\eta_p^2 = 0.407$ (Fig. 4), exhibiting a learning curve that was well accounted for by a linear trend, $F(1,35) = 44.34$, $p < 0.001$, $\eta_p^2 = 0.559$. Once again, performance began at chance level, plateauing around the 35th time each scene was presented. Clearly, the feedback manipulation was effective in modulating the selection behavior of participants.

3.2.2. Test phase

A robust cuing effect of 18 ms, the same magnitude as in Experiment 1, was evident in the test phase in RT, $t(35) = 3.14$, $p = 0.002$, $d = 0.55$ (Fig. 3). Accuracy was once again generally high and did not exhibit a cuing effect (valid = 93.9%, invalid = 93.3%), $t(35) = 1.12$, $p = 0.271$, and there was no hint of a speed-accuracy tradeoff as participants were numerically less accurate on invalid trials.

The results of Experiment 2 are clear: monetary reward is not necessary to observe the learning and expression of habitual spatial attention biases. Rather, both the learning and persistent expression of attentional bias observed in Experiment 1 appear to be driven by reinforcement-guided selection history. Through feedback, participants are encouraged to repeatedly select a particular region of a scene. This repeated selection, which may or may not be driven by internal reward signals (see *7. General Discussion*), becomes automatic such that the mere presence of the scene comes to trigger a spatial attention shift.

This contrasts with attentional biases towards prior target features, which are not typically evident following brief unrewarded training (see, e.g., Anderson, 2016c; Anderson & Halpern, 2017; Anderson, Laurent, & Yantis, 2014; Anderson et al., 2011, 2012; Qi, Zeng, Ding, & Li, 2013; Roper & Vecera, 2016; see also Sali, Anderson, & Yantis, 2014). Biases for stimulus features arising from reward-independent selection history typically require extensive training to develop (e.g., Anderson, Chiu, DiBartolo, & Leal, 2017; Kyllingsbaek, Schneider, & Bundesen, 2001; Kyllingsbaek, Van Lommel, Sorensen, & Bundesen, 2014; Qu, Hillyard, & Ding, 2017; Shiffrin & Schneider, 1977), whereas in the present study each scene was only presented forty times during training. One salient difference in the training protocols used across studies is that for studies examining value-driven attention to stimulus features, the reward learning is only loosely contingent upon performance (must correctly report the target in an easy color-search task in which accuracy is generally high), and is thought to rely predominantly on Pavlovian mechanisms (see esp., Le Pelley et al., 2015; Sali et al., 2014). On the other hand, the training phase of the present study clearly emphasizes reinforcement learning, as reward is entirely contingent on the spatial choices of participants. To the degree that learning in the feature and spatial domain predominantly rely upon different underlying learning mechanisms, their specificity with regards to the nature of performance feedback may be different (see Section 7 for further discussion on this issue).

4. Experiment 3

The findings of Experiment 2, in which non-reward feedback both curbed spatial choice behavior and biased spatial attention in a subsequent task, raises an important question concerning the reason why significant biases were observed in the present study but not in Jiang et al. (2015) and Won and Leber (2016). As previously suggested, one possibility is that the scene context allows participants to anchor spatial representations to the configuration of objects within a scene (e.g., to the left of object A and above object B) or to its unique spatial layout (e.g., bottom-left of room A). Such configural spatial information processing could help explain why Chelazzi et al. (2014) reported a significant spatial attention bias on two-target but not single-target trials. Another possibility is that different learning mechanisms, one Pavlovian and the other reinforcement-based, is alone responsible for the different pattern of results. In this case, any repeated shift of spatial attention can be trained to become automatic within a particular context with sufficient repetition. Note that in Jiang et al. (2015) and Won and Leber (2016), the target appeared equally-often in each quadrant when the effects of reward were examined, such that no one quadrant was associated with a greater frequency of selections. In this sense, the reward coincided with where the target happened to be on a given trial in those studies, and was delivered probabilistically based on this contingency. In the present study, the selection behavior of participants directly determined the amount of reward received.

To test between these two competing accounts, we conducted a third experiment in which scene textures, rather than scenes containing a unique arrangement of meaningful objects and a distinct spatial layout, were used. Such scene textures are more analogous to the uniform background on which letter-like characters appear in studies failing to show spatial attention biases following reward training (Jiang et al., 2015; Won & Leber, 2016). To maximize scene-specific learning, we only trained four rather than eight scenes, reducing the overall memory burden for rewarded locations. The texture scenes contained no objects and were roughly uniform throughout, but were very easily distinguishable from each other, allowing for context-specific learning. The same feedback manipulation was used as in Experiment 1. Of interest was, in the event of robust learning and thus repeated selection of particular regions of different scenes during training, can a subsequent attentional bias still be observed. If the spatial representations that bias attention are anchored to the configuration of objects within a scene or its unique spatial layout, no bias should be evident, but if reinforcing a particular spatial shift of attention in a particular context is alone sufficient to produce a spatial attention bias, the results of Experiment 3 should mirror those of the prior two experiments.

4.1. Methods

4.1.1. Participants

Thirty-six new participants were recruited from the Texas A&M University community. Participants were compensated with money earned in the experimental task. All reported normal or corrected-to-normal visual acuity and normal color vision. Data from one participant was replaced due to chance-level performance in the test phase (accuracy = 46%).

4.1.2. Experimental task

The experimental task was identical to that used in Experiment 1, with the following exceptions. Four texture images were used for scenes. The textures were of sand, water, rock, and an aerial view of a dense forest, and were taken from images freely available on the

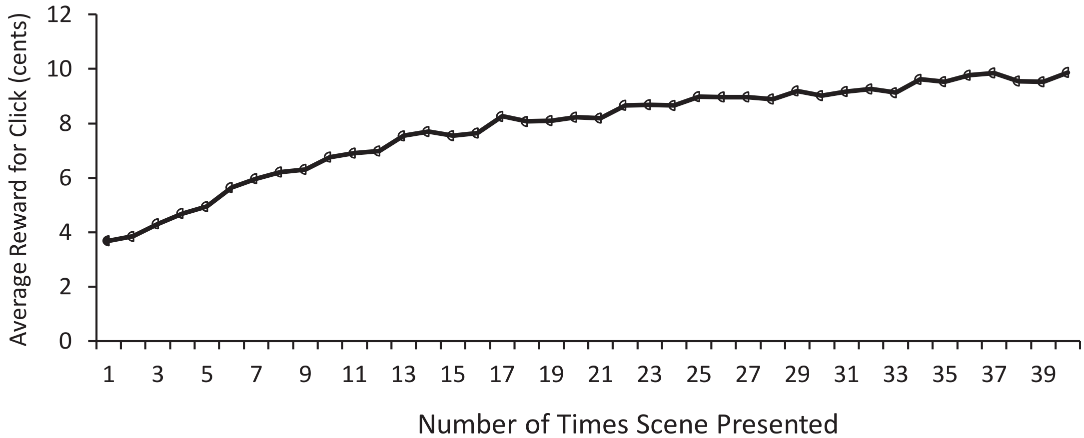

Fig. 5. Reward earned across different presentations of each scene during the training phase of Experiment 3.

internet. During training, each scene was presented on 40 trials, reducing the duration of training to 160 total trials. Each of the four quadrants was the high-value quadrant for one of the scene textures, with each quadrant serving as the high-value quadrant equally-often for each scene texture across participants. Rewards were increased to 2, 8, and 15 cents to provide participants with an appropriate level of overall compensation.

4.2. Results and discussion

4.2.1. Training phase

As in the prior two experiments, performance differed across scene presentation, $F(39,1365) = 44.62$, $p < 0.001$, $\eta_p^2 = 0.560$ (Fig. 5), exhibiting a learning curve that was well accounted for by a linear trend, $F(1,35) = 114.99$, $p < 0.001$, $\eta_p^2 = 0.767$. Even without the presence of objects with which to anchor representations of space, participants had no difficulty learning to repeatedly select higher-value regions of a scene texture.

4.2.2. Test phase

Unlike in the prior two experiments, no cuing effect was evident in either RT, $t(35) = 0.16$, $p = 0.874$ (Fig. 3), or accuracy (valid = 95.6%, invalid = 95.3%), $t(35) = 0.79$, $p = 0.437$. Importantly, the magnitude of the cuing effect was significantly larger in Experiment 1, $t(70) = 2.25$, $p = 0.027$, $d = 0.53$, and Experiment 2, $t(70) = 2.69$, $p = 0.009$, $d = 0.63$, when comparing to the non-significant cuing effect observed here. Even though participants repeatedly selected a particular region of space in each scene texture during training, which was evident in a robust learning curve, this repeated behavior was not effective in creating an enduring attentional bias.

Performance was generally faster in Experiment 3 (mean RT = 761 ms) compared to the other two experiments (mean RT = 816 ms for Experiment 1 and 837 for Experiment 2), $ts > 2.06$, $ps < 0.043$, likely owing to the reduced difficulty of searching against a more uniform background in the case of scene textures. One possibility is that this did not allow sufficient time for the scene to be processed such that a spatial attention bias could exert itself, whereas this was not the case in Experiments 1 and 2. Two sources of evidence argue against this possible explanation for the different pattern of results. First, in Experiment 3, mean RT was a poor predictor of the cuing effect, $r = 0.042$, $p = 0.811$. Second, when comparing participants in Experiment 3 to the fastest half of participants in each of Experiments 1 and 2, a significant difference in the cuing effect still emerged, $t(70) = 2.41$, $p = 0.019$, $d = 0.57$, even though mean RT was now numerically slower for Experiment 3 (761 vs 738 ms).

It is worth noting that the training phase of Experiment 3 was half the duration of Experiments 1 and 2, which was done to match the number of exposures to each individual scene. It is unclear whether this difference had any effect on overall learning. Although learning was generally robust, with a pronounced learning curve evident in performance, participants also had less time to consolidate learning for each scene prior to the test phase (less time, on average, between two presentations of the same scene), and such consolidation may be important for the transfer of learning to the test phase. At the same time, there were fewer scenes to keep track of during learning, which could enhance the quality or fidelity of learning for each scene.

Taken together, the results of Experiment 3 demonstrate that repeatedly selecting a particular region of space, even when participants are provided monetary rewards for doing so, is not itself sufficient to give rise to an enduring spatial attention bias. In this way, our results are very much in accord with those of Jiang et al. (2015) and Won and Leber (2016). Our findings suggest that the ability to anchor spatial representations to the configuration of objects within a naturalistic scene and its distinct spatial layout is an important component of selection history as it pertains to the biasing of spatial attention.

5. Experiment 4

Another source of ambiguity, both in the spatial attention bias observed in Experiments 1 and 2 and also in prior demonstrations of spatial attention biases driven by selection history (Chelazzi et al., 2014), is the degree to which these biases reflect an automatic stimulus-evoked shift of attention on the one hand versus a bias to execute a particular search pattern or strategy on the other hand. This is because, in the test phase, participants are tasked with searching the display for a target, and attention was probed after only a brief presentation of the search context. Under these conditions, if participants choose to begin search in the previously high-value region or if the presentation of the search context triggers a transient shift of attention, a spatial attention bias is predicted in each case. However, if the stimuli that would trigger the shift of attention remain on screen for a sufficient period of time, participants will reorient their attention and perhaps exhibit inhibition of return (IOR; Klein, 2000; Posner, Rafal, Choate, & Vaughan, 1985). This is not the case for the execution of a search pattern or strategy, which would be engaged whenever the search array is onset. Experiment 4 tested between these two accounts by increasing the stimulus-onset-asynchrony (SOA) between the presentation of the scene context and the presentation of the search array (letters) during the test phase.

5.1. Methods

5.1.1. Participants

Thirty-six new participants were recruited from the Texas A&M University community. Participants were compensated with course credit. All reported normal or corrected-to-normal visual acuity and normal color vision. Data from one participant was replaced due to chance-level performance in the test phase (accuracy = 47%).

5.1.2. Experimental task

The experimental task was identical to that used in Experiment 2, with the following exceptions. Only four of the eight scenes were used, each being presented on 40 trials during training for a total of 160 training trials. Each of the four quadrants was the quadrant that yielded positive feedback for one of the scenes, with each quadrant serving as the reinforced quadrant equally-often for each scene across participants. During the test phase, the period of time over which the scene was presented prior to the onset of the search array was increased to 2500 ms.

5.1.3. Data analysis

The data were analyzed in the same manner as Experiment 2, substituting feedback with numerical values for the sake of facilitating statistical analysis.

5.2. Results and discussion

5.2.1. Training phase

As in Experiment 2, performance differed across scene presentation, $F(39,1365) = 22.19$, $p < 0.001$, $\eta_p^2 = 0.388$, exhibiting a learning curve that was well accounted for by a linear trend, $F(1,35) = 74.32$, $p < 0.001$, $\eta_p^2 = 0.680$.

5.2.2. Test phase

Unlike Experiment 2, no cuing effect was evident in either RT, $t(35) = 0.29$, $p = 0.775$ (Fig. 3), or accuracy (valid = 95.4%, invalid = 94.9%), $t(35) = 1.32$, $p = 0.197$. The difference in the cuing effect between this experiment and Experiment 1, $t(70) = 1.00$, $p = 0.321$, and well as Experiment 2, $t(70) = 1.04$, $p = 0.303$, did not reach the threshold for statistical significance. However, across participants, the variability of the cuing effect was substantially greater in this experiment compared to both Experiment 1, $F(35,35) = 3.35$, $p < 0.001$, and Experiment 2, $F(35,35) = 5.49$, $p < 0.001$. The Bayes factor for the cuing effect in Experiment 4 was also 5.30 in favor of the null hypothesis, which is considered strong evidence commensurate with traditional thresholds for rejecting the null using conventional hypothesis testing methods (Rouder, Speckman, Sun, Morey, & Iverson, 2009). The number of participants exhibiting a large negative cuing effect (< −50 ms) indicative of IOR was substantially greater in Experiment 4 than in the prior three experiments combined (9 vs 0), $\chi^2 = 28.80$, $p < 0.001$, $\varphi = 0.447$.

Lengthening the scene-to-search array SOA resulted in the absence of a reliable cuing effect across participants, in contrast to the results of Experiment 2. Interestingly, there appeared to be substantial individual differences in the magnitude of the cuing effect in the present experiment, with some participants exhibiting a large negative cuing effect indicative of IOR (unlike in any of the prior experiments) and others exhibiting a large positive cuing effect indicative of attentional orienting. Perhaps the time course of the processing of the scene semantics necessary to guide spatial attention is variable across individuals, such that only some participants had sufficient time to disengage attention prior to the onset of the search array. Taken together, the results of Experiment 4 are clearly inconsistent with a bias to execute a particular search strategy, which participants must wait to execute until the search array is presented, and are more consistent with a transient, scene-evoked orienting response.

6. Experiment 5

Experiment 5 sought to provide an even stronger test of whether reinforcement learning biases scene-evoked shifts of spatial attention rather than the execution of a specific search strategy. To this end, we had participants complete a free-viewing task during the test phase, and measured eye position as they viewed the familiar scenes. If reinforcement learning results in an enduring bias to direct attention to a particular region of a scene, participants should preferentially direct their gaze towards the previously high-value quadrant, even though they have no specific motivation for doing so. Such a bias, if evident, would be quite powerful, as it would need to be sufficient to overcome any novelty-seeking bias (e.g., Johnston, Hawley, Plewe, Elliott, & DeWitt, 1990; Johnston & Schwarting, 1997) by which participants prefer to explore less familiar aspects of the scene. We looked for both an immediate bias in initial saccades, as well as a sustained bias in overall frequency of saccades and total time spent fixating the high-value quadrant. Although an initial selection bias may be evident, consistent with the results of Experiment 4, unlike Experiment 4 there is no explicit task and thus no explicit goal-directed process to compete with a bias resulting from selection history. As such, Experiment 5 provides a direct test of spatial attention biases that are independent of search strategy.

6.1. Methods

6.1.1. Participants

Thirty-six new participants were recruited from the Texas A&M University community. Participants were compensated with course credit. All reported normal or corrected-to-normal visual acuity and normal color vision.

6.1.2. Experimental task

The training phase was exactly identical to that of Experiment 2. In the test phase, participants completed a free-viewing task. Eye position was monitored throughout the test phase using an EyeLink 1000 Plus eye tracker running at 1000 Hz. The test phase consisted of two blocks of 48 trials, during which each scene was presented six times for 3000 ms each in a random order. Scenes were presented successively without enforced central fixation between scenes. Nine-point calibration was used at the beginning of each block. For a dataset to be considered usable, all points needed to be validated within 1.5°. Participants were informed that there was no specific task to perform, and that they were free to look around the images however they wished.

6.1.3. Data analysis

The training phase data were analyzed in the same manner as Experiment 2, substituting feedback with numerical values for the sake of facilitating statistical analysis. For the test phase, we compared two indices of attentional bias between the previously high-value quadrant and the average of the three previously low-value quadrants. The first was total fixation time, summed over all fixations within each quadrant, and the second was saccade shifts to each quadrant. In defining saccade shifts and fixation time, four areas of interest (AOIs) were drawn, each encompassing an entire quadrant of the screen. A saccade shift was defined as when a saccade crossed one of the AOI boundaries, landing in a different quadrant than where it originated. In a follow-up analysis examining the first saccade shift, only trials on which eye position did not begin in the high-value quadrant were analyzed, and the percentage of saccade shifts to the high-value quadrant was measured for each participant.

6.2. Results and discussion

6.2.1. Training phase

As in Experiment 2, performance differed across scene presentation, $F(39,1365) = 59.23$, $p < 0.001$, $\eta_p^2 = 0.629$, exhibiting a learning curve that was well accounted for by a linear trend, $F(1,35) = 116.54$, $p < 0.001$, $\eta_p^2 = 0.769$.

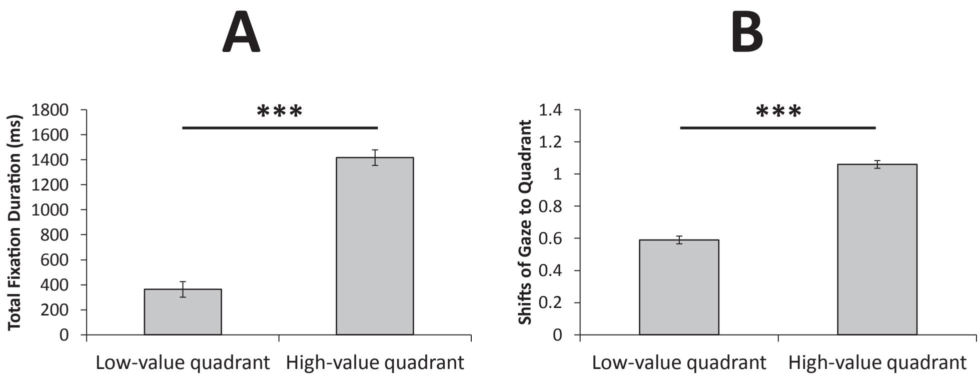

Fig. 6. Eye position data from Experiment 5. (A) Total fixation duration, summed over all fixations within a quadrant, for the high-value compared to the average of the low-value quadrants, averaged over each scene presentation. (B) Number of saccades made to the high-value compared to the average of the low-value quadrants, averaged over each scene presentation.

6.2.2. Test phase

Eye movements were robustly biased towards the previously high-value quadrant. Participants spent a longer amount of time fixating the previously high-value quadrant compared to the other three quadrants, $t(35) = 8.47, p < 0.001, d = 1.41$. Participants also shifted gaze more frequently to the previously high-value quadrant compared to the other three quadrants, $t(35) = 9.90, p < 0.001, d = 1.65$ (see Fig. 6). When eye position began in a low-value quadrant, 64.2% of first saccade shifts were made to the high-value quadrant, which was significantly more than would be expected from unbiased selection (33.3%), $t(35) = 10.22, p < 0.001, d = 1.70$. Even though there was no specific task or motivation to look towards any one quadrant more than another, participants preferentially looked at the quadrant that they had more frequently selected during the training phase, providing compelling evidence that reinforcement learning is capable of generating an intrinsic attentional bias towards a particular region within a scene.

7. General discussion

Although the ability of reward learning to modulate feature-based attention is well established (see Anderson, 2016b), whether reward learning can influence spatial attention has been much more controversial (Chelazzi et al., 2014; Jiang et al., 2015; Won & Leber, 2016). In the present study, we sought to characterize whether and how selection history influences spatial attention orienting. Our findings offer new insights into the mechanisms and principles governing the interplay between these two systems, providing an integrative account that bridges previously conflicting findings in the literature.

7.1. Representational basis

When scenes contained an array of objects with a meaningful spatial organization, reinforcing selection of a particular region within a given scene resulted in an enduring spatial attention bias towards that region. This provides strong evidence that learning from feedback can robustly shape the spatial attention system, consistent with an earlier report (Chelazzi et al., 2014). A spatial bias was not evident, however, when scenes provided an easily identifiable spatial context, but were largely uniform in their spatial content (scene textures), without any objects or boundaries by which to anchor position information (e.g., to the left of the window and above the bed). The absence of a learned spatial bias under these conditions echoes more recent reports demonstrating negligible influences of reward feedback on spatial attention (Jiang et al., 2015; Won & Leber, 2016).

Our findings suggest that relational information within a scene is an important component of the underlying representation that is modified by selection history to guide spatial attention. One possibility is that observers anchored their representation of space to the configuration of meaningful objects within the scenes. In naturalistic environments, landmarks often serve as the basis for spatial orienting. Humans rely strongly on landmarks when navigating (e.g., Montello, 2005; Newman et al., 2006). The fewer the landmarks that are available to assist with navigation, and the more individuals have to rely on body-centered representations involved in dead reckoning, the poorer navigation performance becomes (Cornell & Heth, 2000). A second possibility is that the spatial layout provided by the scenes, which depict clear boundaries not found in the scene textures, served as the anchor point for spatial attention (e.g., lower-left side of the room). In the absence of landmarks and layout information, a condition represented by scene textures in the present study, spatial attention does not seem to be sensitive to influences of selection history. In this sense, the spatial attention biases observed in the present study map well onto the spatial information that humans use to orient and navigate in naturalistic environments, supporting the ecological validity of this automatic orienting mechanism.

In this way, the findings lend important insights into the discrepancy between the positive findings of Chelazzi et al. (2014) on the one hand and the negative findings of Jiang et al. (2015) and Won and Leber (2016) on the other hand. The displays used by Jiang et al. (2015) and Won and Leber (2016) are clearly more similar to the scene textures used in the present study, where there is no clear anchor point for the spatial representations upon which reward is contingent. Let us assume that the representation of space utilized by the participants in these studies covers the extent of the computer screen (rather than the entire room in which the experiment is conducted). Rewards are greater for targets appearing in a particular quadrant of this space (e.g., upper right). There is nothing intrinsic about the space that would suggest to participants that it should be partitioned in this way. It is possible that participants would divide the extent of the spatial layout in two in each direction, thereby determining the mid-point, and then use the mid-point to partition the space. An infinite number of other arbitrary partitions of the space are also possible, or no partition at all (i.e., represent the space as a whole). The main alternative would be to use a purely body-centered representation to define the space, which does not seem to be the case in either these prior studies or in the present study. Again, this fits with how observers tend to use spatial information to orient and navigate (e.g., Cornell & Heth, 2000; Montello, 2005; Newman et al., 2006).

Interestingly, Chelazzi et al. (2014) only found evidence for a learned spatial attention bias when two targets were presented simultaneously in a data-limited (i.e., briefly presented) search array, but not when a single target was presented. These authors interpreted this difference in terms of the increased difficulty of two-target trials providing greater sensitivity to learning effects, although performance was well below ceiling on single-target trials. Another possibility, not mutually exclusive with the interpretation forwarded by the authors, is that on two-target trials, participants represented the configuration of the two high-value targets, or the position of one relative to the other. To the degree that participants represented the displays in this sort of way, and it is this representation in particular that is shaped by selection history, the proposed mechanism can explain the apparent discrepancy between Chelazzi et al. (2014) and Jiang et al. (2015) and Won and Leber (2016), in addition to the specific pattern of results observed by Chelazzi et al. (2014).

The spatial representations that have been argued to be sensitive to learned biases arising from selection history are not “purely spatial” insofar as they rely in part on object processing mechanisms. This is in contrast to the representations of space probed by the scene textures and in prior studies employing a uniform search context (Jiang et al., 2015; Won & Leber, 2016). It is important to note that such “impure” representations of space are of the sort that are typically investigated in the field of “spatial” cognition and thought to play a central role human navigation ability (e.g., Montello, 2005; Newman et al., 2006). It is difficult to think that humans could develop a robust spatial attention bias within an environment that they would be largely unable to navigate. It is also debatable whether object-independent representations of space are any less “pure” in their spatial nature, as they must still be defined in relation to something—namely body position.

Nevertheless, it could be argued that there is nothing intrinsically spatial at all about the representations used to guide attention in the present study. Perhaps participants selectively attend to a particular object within the high-value quadrant, and the observed cuing effects merely reflect proximity to this reward-associated object. This is difficult to rule out in practice, as space and the objects that define it are intricately intertwined. The scenes used in the present study tended to include a rich array of multiple objects from the same category, as well as objects that spanned multiple quadrants. A number of different types of objects appeared within multiple difference scenes as well (e.g., books on a bookshelf, chairs, beds, cabinets, pictures on a wall, chests of drawers, windows, etc.). It would appear that simple object-based biases, which tend to show appreciable tolerance within a particular object category (e.g., Hickey, Keiser, & Peelen, 2015; Hickey & Peelen, 2015), cannot easily account for the findings of the present study. Furthermore, the ability to represent object identity presumably requires spatial attention directed to the object (e.g., Treisman & Gelade, 1980). Perhaps the most telling evidence that object-based attentional biases cannot provide a complete account of the present findings can be found in the fundamental difference in the learning mechanisms that are involved (see Section 7.3 below).

Also speaking to the ecological validity of the observed spatial attention biases, these biases were by definition contextually specific. It was not the case that one region of space (e.g., upper right side of the screen) was any more rewarded than any other region across scenes, or that any one type of object was itself predictive of reward. Rather, the observed spatial attention biases were specific to the particular arrangement of a particular set of objects, with multiple such arrangements stored in long-term memory to guide selection. In this way, spatial attention biases are sensitive to context in a similar manner to what has been shown to be the case for feature-based attentional biases (Anderson, 2015b), allowing for automatic attention influences to more reliably guide selection than would otherwise be the case.

7.2. Mechanisms of attentional selection

The findings of the present study suggest that selection history can bias stimulus-evoked spatial orienting. The observed attentional bias was shown to be transient, no longer evident if the scene context was presented sufficiently in advance of the search array, which is inconsistent with the bias reflecting the execution of a particular search strategy (e.g., search and disconfirm the upper right quadrant before considering other areas). A subset of participants exhibited pronounced IOR, further consistent with stimulus-evoked orienting. Compellingly, eye movements were robustly biased towards the previously high-value quadrant during a free-viewing task.

The observed spatial attention bias also appears to be non-strategic. There was no benefit to orienting to the previously reinforced region of a scene during the test phase; indeed, doing so would direct participants to a non-target location more often than not in Experiment 1–4, and Experiment 5 had no specific task. Similar logic is used in studies of contingent attentional capture that employ the spatial cuing paradigm (e.g., Folk & Remington, 1998; Folk, Remington, & Johnston, 1992). Taken together, the results of the present study suggest that when observers are confronted with a familiar scene in which a spatial shift of attention to a particular region has been reinforced, this spatial shift of attention comes to be evoked by the scene it is associated with.

7.3. Learning mechanisms

Unlike value-driven attentional capture by previously rewarded features, which converging evidence suggests is learned via Pavlovian rather than reinforcement learning mechanisms (e.g., Bucker & Theeuwes, 2017; Le Pelley et al., 2015; Sali et al., 2014), the influence of selection history on spatial attention appears to be distinctly reinforcement-based. Extrinsic (monetary) reward is clearly not needed to observe a robust effect of selection history, as non-reward corrective feedback produced a bias of identical magnitude. This contrasts sharply with studies examining the effects of unrewarded selection history on feature-based attention, which frequently produce no evidence of a persistent attentional bias (e.g., Anderson, 2016c; Anderson & Halpern, 2017; Anderson et al., 2011, 2012; Anderson, Laurent, et al., 2014; Qi et al., 2013; Roper & Vecera, 2016; see also Sali et al., 2014). This discrepancy cannot be readily explained by simply appealing to the strength of the training manipulation, as these feature-based attention studies included many more trials per trained stimulus than was used in the present study (i.e., 40). Value-dependent attentional biases have also been demonstrated for differently rewarded objects, faces, and scenes matched for selection history (Barbaro, Peelen, & Hickey, 2017; Della Libera & Chelazzi, 2009; Donohue et al., 2016; Failing & Theeuwes, 2015; Hickey & Peelen, 2015; Kim, Ghazizadeh, & Hikosaka, 2015; Raymond & O'Brien, 2009; Yamamoto, Kim, & Hikosaka, 2013), which contrasts with the findings of the present study in which the magnitude of attentional bias was unaffected by the presence or absence of associated extrinsic rewards (compare Experiment 1 and 2).

We propose that the effects of selection history on spatial attention operate over fundamentally different learning mechanisms than those thought to underlie value-driven attention to high-value features and objects. Participants execute a spatial shift of attention, very likely both covert and overt in nature (given the task to move a cursor to the selected location), to a chosen region of space. This spatial shift of attention can be thought of as a reinforceable behavior, just like a motor response (e.g., pressing a lever). Based on the quality of the resulting feedback, participants are either inclined to repeat this behavior or execute a different behavior the next time the scene is encountered. As a particular shift of attention is repeated in a particular context, it eventually becomes automatic in that context such that the presentation of the context itself triggers the associated behavior.

An open question remains the extent to which the reward system influences the spatial attention system, and whether the reinforcement learning observed in the present study is directly or indirectly reward-related. One possibility is that the non-reward corrective feedback used in Experiments 2 and 5 motivates participants to repeat a particular behavior, and this repetition of behavior is alone sufficient to produce a spatial attention bias. Should this be the case, simply instructing participants to select a particular region of a scene, rather than having them discover the "best" location to select through trial and error, would be sufficient to produce a robust attentional bias. Another possibility is that the non-reward feedback, when positive, resulted in internal reward signals, as has been hypothesized to play a role in perceptual learning (Herzog & Fahle, 1999; Roelfsema & van Ooyen, 2005; Roelfsema, van Ooyen, & Watanabe, 2010; Sasaki, Nanez, & Watanabe, 2010; Seitz, Lefebvre, Watanabe, & Jolicoeur, 2005; Seitz & Watanabe, 2005). In fact, even the simple act of correctly completing a trial has been hypothesized to produce an internal reward signal capable of modulating sensory representations through a learning process (Sasaki et al., 2010; Seitz & Watanabe, 2005). Teasing these two competing possibilities apart is difficult, as imposing any task structure that results in the ability to produce correct vs incorrect behavior would have the potential to generate associated internal reward signals. One potential approach would be to examine spatial attentional learning in individuals who differ in how they process reward information, for example individuals who are depressed (see Anderson, Leal, Hall, Yassa, & Yantis, 2014; Anderson et al., 2017).

A related question concerns the role of awareness in the observed spatial attention biases. Explicit awareness of the reward contingencies does not appear to be critical for attentional biases towards previously reward-associated features (e.g., Anderson, 2015a, 2015b; Anderson, Faulkner, Rilee, Yantis, & Marvel, 2013; Bourgeois, Neveu, & Vuilleumier, 2016; Leganes-Fonteneau, Scott, & Duka, 2018; Pearson, Donkin, Tran, Most, & Le Pelley, 2015), nor is explicit awareness of repeated stimulus configurations necessary for contextual cuing (Chun & Jiang, 1998, 2003). As such, implicit learning mechanisms may reflect a core component of the influence of selection history on attention, including spatial attention. At the same time, we argue that the attentional biases observed in the present study are the consequence of reinforcement learning, which contrasts with attentional biases towards previously reward-associated features, which are presumed to operate via Pavlovian learning mechanisms (e.g., Bucker & Theeuwes, 2017; Le Pelley et al., 2015; Sali et al., 2014). The overt choice component of the training phase of the present study encourages active learning through trial and error, and such active learning may be critical for the attentional biases we observed, which might explain the equally robust bias with and without explicit (monetary) reward. Consistent with this distinction, explicit knowledge has been implicated in scene-based contextual cueing (Brockmole & Henderson, 2006b). Future research might employ a training procedure in which reward is decoupled from the decision where to orient/click in order to directly examine the role of contingency awareness in biasing spatial attention.

7.4. Conclusions

The findings of the present study highlight a powerful role for reinforcement learning in the control of spatial attention. This influence does not depend specifically on reward feedback, but rather reflects a broader consequence of selection history. When observers repeat a spatial shift of attention to a particular region of a scene in order to achieve a desired outcome, the scene will come to automatically trigger this attention shift. This learned spatial bias is tied to relational information contained within a particular scene (either the configuration of objects, the spatial layout, or both), and can be maintained for a variety of different scenes with context specificity. In this way, rewarding outcomes can bias attention in at least two fundamentally different ways, one on the basis of reward-predictive features and one on the basis of spatially-localized shifts of attention that are reinforced by the resulting outcome. The influence of reinforcement learning on spatial attention reflects a new vista in attention research that is likely to offer important insights into how different aspects of selection history (Awh et al., 2012) differently influence the control of information processing.

Funding statement

The reported research was funded through a start-up package to BAA through Texas A&M University and NARSAD Young Investigator Grant 26008 to BAA.

Appendix A. Supplementary material

Raw data for the training and test phases of all experiments. Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2018.05.005.

References

Anderson, B. A. (2013). A value-driven mechanism of attentional selection. *Journal of Vision, 13*(3:7), 1–16.
Anderson, B. A. (2015a). Value-driven attentional capture is modulated by spatial context. *Visual Cognition, 23*, 67–81.
Anderson, B. A. (2015b). Value-driven attentional priority is context specific. *Psychonomic Bulletin and Review, 22*, 750–756.
Anderson, B. A. (2016a). Social reward shapes attentional biases. *Cognitive Neuroscience, 7*, 30–36.
Anderson, B. A. (2016b). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences, 1369*, 24–39.
Anderson, B. A. (2016c). Value-driven attentional capture in the auditory domain. *Attention, Perception, and Psychophysics, 78*, 242–250.
Anderson, B. A., Chiu, M., DiBartolo, M. M., & Leal, S. L. (2017). On the distinction between value-driven attention and selection history: Evidence from individuals with depressive symptoms. *Psychonomic Bulletin and Review, 24*, 1636–1642.
Anderson, B. A., Faulkner, M. L., Rilee, J. J., Yantis, S., & Marvel, C. L. (2013). Attentional bias for non-drug reward is magnified in addiction. *Experimental and Clinical Psychopharmacology, 21*, 499–506.
Anderson, B. A., Folk, C. L., Garrison, R., & Rogers, L. (2016). Mechanisms of habitual approach: Failure to suppress irrelevant responses evoked by previously reward-associated stimuli. *Journal of Experimental Psychology: General, 145*, 796–805.
Anderson, B. A., & Halpern, M. (2017). On the value-dependence of value-driven attentional capture. *Attention, Perception, and Psychophysics, 79*, 1001–1011.
Anderson, B. A., Kuwabara, H., Wong, D. F., Gean, E. G., Rahmim, A., Brasic, J. R., ... Yantis, S. (2016). The role of dopamine in value-based attentional orienting. *Current Biology, 26*, 550–555.
Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences, USA, 108*, 10367–10371.
Anderson, B. A., Laurent, P. A., & Yantis, S. (2012). Generalization of value-based attentional priority. *Visual Cognition, 20*, 647–658.
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.
Anderson, B. A., Leal, S. L., Hall, M. G., Yassa, M. A., & Yantis, S. (2014). The attribution of value-based attentional priority in individuals with depressive symptoms. *Cognitive, Affective, and Behavioral Neuroscience, 14*, 1221–1227.
Anderson, B. A., & Yantis, S. (2012). Value-driven attentional and oculomotor capture during goal-directed, unconstrained viewing. *Attention, Perception, and Psychophysics, 74*, 1644–1653.
Anderson, B. A., & Yantis, S. (2013). Persistence of value-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 6–9.
Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*, 437–443.
Barbaro, L., Peelen, M. V., & Hickey, C. (2017). Valence, not utility, underlies reward-driven prioritization in human vision. *Journal of Neuroscience, 37*, 10438–10450.
Bourgeois, A., Neveu, R., & Vuilleumier, P. (2016). How does awareness modulate goal-directed and stimulus-driven shifts of attention triggered by value learning? *PLoS ONE, 11*, e0160469.
Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.
Brockmole, J. R., & Henderson, J. M. (2006a). Recognition and attentional guidance during contextual cueing in real-world scenes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology, 59*, 1177–1187.
Brockmole, J. R., & Henderson, J. M. (2006b). Using real-world scenes as contextual cues for search. *Visual Cognition, 13*, 99–108.
Bucker, B., & Theeuwes, J. (2017). Pavlovian reward learning underlies value driven attentional capture. *Attention, Perception, and Psychophysics, 79*, 415–428.
Chelazzi, L., Estocinova, J., Calletti, R., Lo Gerfo, E., Sani, I., Della Libera, C., & Santandrea, E. (2014). Altering spatial priority maps via reward-based learning. *Journal of Neuroscience, 34*, 8594–8604.
Chun, M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28–71.
Chun, M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 224–234.
Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*, 201–215.
Cornell, E. H., & Heth, C. D. (2000). Route learning and wayfinding. In R. Kitchin, & S. Freundschuh (Eds.). *Cognitive mapping: Past, present, and future* (pp. 66–83). New

York: Routledge.
Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science, 20*, 778–784.
Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.
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, 529–541.
Failing, M. F., & Theeuwes, J. (2014). Exogenous visual orienting by reward. *Journal of Vision, 14*(5:6), 1–9.
Failing, M. F., & Theeuwes, J. (2015). Nonspatial attentional capture by previously rewarded scene semantics. *Visual Cognition, 23*, 82–104.
Folk, C. L., & Remington, R. W. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 847–858.
Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 1030–1044.
Herzog, M. H., & Fahle, M. (1999). Effects of biased feedback on learning and deciding in a vernier discrimination task. *Vision Research, 39*, 4232–4243.
Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience, 30*, 11096–11103.
Hickey, C., Chelazzi, L., & Theeuwes, J. (2014). Reward priming of location in visual search. *PLoS ONE, 9*(7), e103372.
Hickey, C., Keiser, D., & Peelen, M. V. (2015). Reward guides attention to object categories in real-world scenes. *Journal of Experimental Psychology: General, 144*, 264–273.
Hickey, C., & Peelen, M. V. (2015). Neural mechanisms of incentive salience in naturalistic human vision. *Neuron, 85*, 512–518.
Jiang, V. Y., Li, Z. S., & Remington, R. W. (2015). Modulation of spatial attention by goals, statistical learning, and monetary reward. *Attention, Perception, and Psychophysics, 77*, 2189–2206.
Jiang, Y. V., & Swallow, K. M. (2013). Spatial reference frame of incidentally learned attention. *Cognition, 126*, 378–390.
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*, 87–99.
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, 119*, 397–411.
Johnston, W. A., & Schwarting, I. S. (1997). Novel popout: An enigma for conventional theories of attention. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 622–631.
Kim, H. F., Ghazizadeh, A., & Hikosaka, O. (2015). Dopamine neurons encoding long-term memory of object value for habitual behavior. *Cell, 163*, 1165–1175.
Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4*, 138–147.
Kyllingsbaek, S., Schneider, W. X., & Bundesen, C. (2001). Automatic attraction of attention to former targets in visual displays of letters. *Perception and Psychophysics, 63*, 85–98.
Kyllingsbaek, S., Van Lommel, S., Sorensen, T. A., & Bundesen, C. (2014). Automatic attraction of visual attention by supraletter features of former target strings. *Frontiers in Psychology, 5*(1383), 1–7.
Laurent, P. A., Hall, M. G., Anderson, B. A., & Yantis, S. (2015). Valuable orientations capture attention. *Visual Cognition, 23*, 133–146.
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*, 158–171.
Lee, J., & Shomstein, S. (2013). Differential effects of reward on space- and object-based attentional allocation. *Journal of Neuroscience, 33*, 10625–10633.
Lee, J., & Shomstein, S. (2014). Reward-based transfer from bottom-up to top-down search tasks. *Psychological Science, 25*, 466–475.
Leganes-Fonteneau, M., Scott, R., & Duka, T. (2018). Attentional responses to stimuli associated with a reward can occur in the absence of knowledge of their predictive values. *Behavioral Brain Research, 341*, 26–36.
Mine, C., & Saiki, J. (2015). Task-irrelevant stimulus-reward association induces value-driven attentional capture. *Attention, Perception, and Psychophysics, 77*, 1896–1907.
Moher, J., Anderson, B. A., & Song, J.-H. (2015). Dissociable effects of salience on attention and goal-directed action. *Current Biology, 25*, 2040–2046.
Montello, D. R. (2005). Navigation. In P. Shah, & A. Miyake (Eds.). *The Cambridge handbook of visuospatial thinking* (pp. 257–294). New York: Cambridge University Press.
Newman, E. L., Caplan, J. B., Kirschen, M. P., Korolev, I. O., Sekuler, R., & Kahana, M. J. (2006). Learning your way around town: How virtual taxicab drivers learn to use both layout and landmark information. *Cognition, 104*, 231–253.
Pearson, D., Donkin, C., Tran, S. C., Most, S. B., & Le Pelley, M. E. (2015). Cognitive control and counterproductive oculomotor capture by reward-related stimuli. *Visual Cognition, 23*, 41–66.
Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2014). Where is the chocolate? Rapid spatial orienting toward stimuli associated with primary reward. *Cognition, 130*, 348–359.
Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology, 2*, 211–228.
Qi, S., Zeng, Q., Ding, C., & Li, H. (2013). Neural correlates of reward-driven attentional capture in visual search. *Brain Research, 1532*, 32–43.
Qu, Z., Hillyard, S. A., & Ding, Y. (2017). Perceptual learning induces persistent attentional capture by nonsalient shapes. *Cerebral Cortex, 27*, 1512–1523.
Raymond, J. E., & O'Brien, J. L. (2009). Selective visual attention and motivation: The consequences of value learning in an attentional blink task. *Psychological Science, 20*, 981–988.
Roelfsema, P. R., & van Ooyen, A. (2005). Attention-gated reinforcement learning of internal representations for classification. *Neural Computation, 17*, 2176–2214.
Roelfsema, P. R., van Ooyen, A., & Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences, 14*, 64–71.
Roper, Z. J. J., & Vecera, S. P. (2016). Funny money: The attentional role of monetary feedback detached from expected value. *Attention, Perception, and Psychophysics, 78*, 2199–2212.
Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review, 16*, 225–237.
Sali, A. W., Anderson, B. A., & Yantis, S. (2014). The role of reward prediction in the control of attention. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 1654–1664.
Sareen, P., Ehinger, K. A., & Wolfe, J. M. (2016). CB Database: A change blindness database for objects in natural indoor scenes. *Behavior Research Methods, 48*, 1343–1348.
Sasaki, Y., Nanez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience, 11*, 53–60.
Seitz, A., Lefebvre, C., Watanabe, T., & Jolicoeur, P. (2005). Requirement for high-level processing in subliminal learning. *Current Biology, 15*, R753–R755.
Seitz, A., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences, 9*, 329–334.
Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing II: Perceptual learning, automatic attending, and general theory. *Psychological Review, 84*, 127–190.
Shomstein, S., & Johnson, J. (2013). Shaping attention with reward: Effects of reward on space- and object-based selection. *Psychological Science, 24*, 2369–2378.
Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica, 135*, 77–99.
Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.
Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided Search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 419–433.
Won, B.-Y., & Leber, A. B. (2016). How do magnitude and frequency of monetary reward guide visual search. *Attention, Perception, and Psychophysics, 78*, 1221–1231.
Yamamoto, S., Kim, H. F., & Hikosaka, O. (2013). Reward value-contingent changes in visual responses in the primate caudate tail associated with a visuomotor skill. *Journal of Neuroscience, 33*, 11227–11238.