OXFORD

Reward processing in the value-driven attention network: reward signals tracking cue identity and location

Brian A. Anderson

Department of Psychology, Texas A&M University, 4235 TAMU, College Station, TX 77843-4235, USA

To whom correspondence should be addressed to Brian A. Anderson, Department of Psychology, Texas A&M University, 4235 TAMU, College Station, TX 77843-4235, USA. E-mail: brian.anderson@tamu.edu

Abstract

Through associative reward learning, arbitrary cues acquire the ability to automatically capture visual attention. Previous studies have examined the neural correlates of value-driven attentional orienting, revealing elevated activity within a network of brain regions encompassing the visual corticostriatal loop [caudate tail, lateral occipital complex (LOC) and early visual cortex] and intraparietal sulcus (IPS). Such attentional priority signals raise a broader question concerning how visual signals are combined with reward signals during learning to create a representation that is sensitive to the confluence of the two. This study examines reward signals during the cued reward training phase commonly used to generate value-driven attentional biases. High, compared with low, reward feedback preferentially activated the value-driven attention network, in addition to regions typically implicated in reward processing. Further examination of these reward signals within the visual system revealed information about the identity of the preceding cue in the caudate tail and LOC, and information about the location of the preceding cue in IPS, while early visual cortex represented both location and identity. The results reveal teaching signals within the value-driven attention network during associative reward learning, and further suggest functional specialization within different regions of this network during the acquisition of an integrated representation of stimulus value.

Key words: selective attention; vision; attention networks; reward learning; fMRI

When a stimulus is learned to predict a reward, that stimulus acquires the ability to automatically capture visual attention even when task-irrelevant and physically inconspicuous (Anderson *et al.*, 2011b). That such attentional orienting is attributable specifically to reward history suggests a distinctly value-driven computation of attentional priority (see Anderson, 2013, for a review). More recent studies have begun to explore the neural mechanisms by which attention is directed to learned predictors of reward.

What has arisen from this emerging literature is a consistent set of findings implicating what I will refer to here as a 'value-driven attention network'; that is, a network of brain regions involved in the signaling of value-based attentional priority. These regions include the caudate tail (Hikosaka *et al.*, 2013; Yamamoto *et al.*, 2013; Anderson *et al.*, 2014; see also Anderson

et al., 2016), object-selective visual cortex/lateral occipital complex (LOC; Anderson *et al.*, 2014; Hickey and Peelen, 2015; Donohue *et al.*, 2016; see also Hopf *et al.*, 2015), and the intraparietal sulcus (IPS; Peck *et al.*, 2009; Qi *et al.*, 2013; Anderson *et al.*, 2014), in addition to early visual cortex (MacLean and Giesbrecht, 2015; van Koningsbruggen *et al.*, 2016; see also Seitz *et al.*, 2009). When previously reward-associated stimuli are processed by the visual system, elevated activity within this network is consistently observed. Whereas the IPS is a part of the dorsal attention network (Corbetta and Shulman, 2002; Corbetta *et al.*, 2008) and contains a spatially organized saliency map (e.g. Bisley and Goldberg, 2003; Balan and Gottlieb, 2006), the caudate tail and LOC comprise the visual corticostriatal loop (Seger, 2013) and robustly represent object identity (e.g. Gill-Spector *et al.*, 2001; Yamamoto *et al.*, 2012). Both receive

input from early visual cortex (e.g. Corbetta and Shulman, 2002; Seger, 2013).

Such priority signals in the value-driven attention network reflect the confluence of visual information (i.e. features of a cue) and (learned) value information, which raises a broader and important question concerning the nature of the information processing that could support their development. The manner in which visual signals are combined with reward signals during learning in attention-related brain areas remains unknown. Some form of reward-modulated visual plasticity would seem to be involved, as predicted by attention-gated reinforcement learning models (Roelfsema and van Ooyen, 2005; Rombouts *et al.*, 2015). It is known that reward signals can evoke prediction-like responses in early visual cortex (Shuler and Bear, 2006; Arsenault *et al.*, 2013), potentially mediating such plasticity. However, the stimulus specificity of reward signals in visual areas has not been explored, nor has the breadth of such signals throughout a broader attention network, beyond early visual cortex.

This study explores the hypothesis that, when a reward is received, the visual representation of the cue that predicted the reward is re-instantiated within the value-driven attention network, thus producing a value-dependent visual signal that could support reward-modulated plasticity. This hypothesis makes two key predictions. First, the receipt of reward should be accompanied by elevated activity within the value-driven attention network, consistent with a teaching signal (Shuler and Bear, 2006; Arsenault *et al.*, 2013). Second and most critically, such elevated activity should contain information about the specific visual cue that preceded the reward, including its spatial position and identity. The first prediction was evaluated using an existing dataset (Anderson *et al.*, 2014), the results of which were then independently replicated using a newly acquired dataset. Regions commonly activated by reward feedback across datasets included the entire value-driven attention network bilaterally. Subsequent analyses evaluated the veracity of the second prediction by exploring the stimulus specificity of the reward signals observed within this network.

Methods

Experiment 1A

Participants. Eighteen neurologically healthy adult volunteers with normal or corrected-to-normal visual acuity and color vision were recruited from the Johns Hopkins University community to participate, as described in Anderson *et al.* (2014). Written informed consent was obtained for each participant. All procedures were approved by the Johns Hopkins Medicine Institutional Review Board.

Behavioral task and procedure. The participant completed 5 runs of a cued reward learning task consisting of 60 trials each (see Figure 1). The experiment utilized a rapid event-related design. Each trial began with a fixation display for 2000 ms, which was followed by a search array for 1000 ms and later by a reward feedback display for 1500 ms. Participants were instructed to search for a target circle that was unpredictably red or green and report the orientation of a bar within the target as either vertical or horizontal via a button press. Half of the trials in each run contained a red target and half contained a green target; each target color appeared on each side of the screen (left or right) equally often. The colors of the non-targets were drawn from the set (blue, cyan, pink, orange, yellow, white) without replacement (as in, e.g. Anderson *et al.*, 2011b, 2014); the
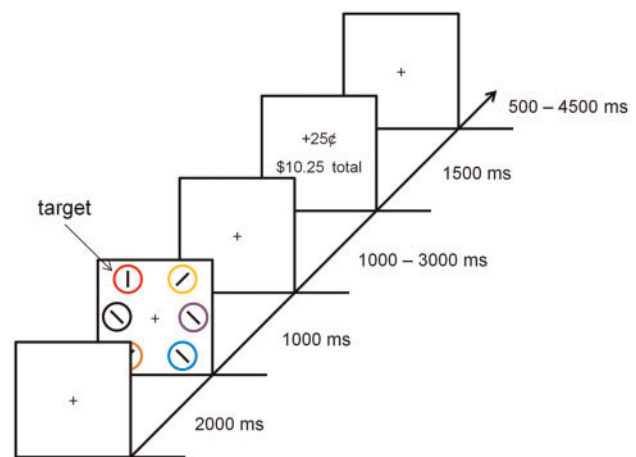

Fig. 1. Sequence and time course of the events for a trial. Participants searched for a color-defined target (red or green) and reported the orientation of the bar within the target as vertical or horizontal. Correct responses resulted in a small amount of money added to the participant's bank total. One color target yielded a high reward for correct responses on 80% of trials (high-reward target), whereas the other color target yielded a high reward on only 20% of correct response trials (low-reward target).

luminance of the colors was not controlled. The order of trials was randomized for each run.

Following a correct response that fell within a 1000 ms response deadline, a small amount of money was added to a bank total in the reward-feedback display. If participants responded incorrectly or too slowly (both were scored as errors), the reward feedback display indicated that 0¢ had been earned for that trial. One of the two target colors was followed by a high reward of 25¢ on 80% of the trials on which it was correctly reported, and by a low reward of 5¢ on the remaining 20% of correct trials (high-reward color); for the other (low-reward) color, these mappings were reversed. The high-reward color was red for half of the participants and green for the other half. An interval during which only the fixation cross was visible was presented between the search array and the reward feedback display for either 1000 or 3000 ms (equally-often), and again immediately following the reward feedback display for 500, 2500 or 4500 ms (exponentially distributed); the fixation cross disappeared for the last 200 ms of the second interval to indicate to the participant that the next trial was about to begin.

Each circle in the search array was 3.4° × 3.4° visual angle in size. The middle of the three circles on each side of the screen was presented 10° center-to-center from fixation, and the two outer circles were presented 8° from the vertical meridian, 6° above and below the horizontal meridian. Participants pressed a button held in the right hand for horizontal targets and a button held in the left hand for vertical targets.

The stimuli were displayed using an Epson PowerLite 7600p projector with a custom zoom lens onto a screen mounted at the end of the magnet bore behind the participant's head. Participants viewed the screen using a mirror mounted to the head coil. Stimulus displays were generated using Matlab software with Psychophysics Toolbox extensions (Brainard, 1997), and responses were recorded using two custom-built, fiber-optic push button boxes.

Each participant practiced the task (without reward feedback) and was trained to a performance criterion (accuracy: 85%, mean RT: 750 ms). Fixation on the central cross was emphasized at all times both during practice and during the

experiment. Data from one run of the training phase was lost for one participant due to equipment failure (computer crash).

MRI data acquisition. Images were acquired using a 3-Tesla Philips Gyroscan MRI scanner and a 32-channel transmit/receive sensitivity encoding (SENSE) head coil at the F. M. Kirby Research Center for Functional Brain Imaging located in the Kennedy Krieger Institute, Baltimore, MD. High-resolution whole-brain anatomical images were acquired using a T1-weighted magnetization-prepared rapid gradient echo pulse sequence [voxel size = 1 mm isotropic, repetition time (TR) = 8.1 ms, echo time (TE) = 3.7 ms, flip angle = 8°, acquisition matrix = 212 × 172, 150 axial slices, 0 mm gap, SENSE factor = 2]. Whole-brain functional images were acquired using a T2*-weighted echoplanar imaging (EPI) pulse sequence (voxel size = 2.5 mm isotropic, TR = 2000 ms, TE = 30 ms, flip angle = 70°, acquisition matrix = 76 × 76, 36 axial slices, 0.5 mm gap, SENSE factor = 2). Each EPI pulse sequence began with four dummy pulses that were not recorded in order allow magnetization to reach steady-state. Each of 5 runs lasted 8.2 min during which 242 volumes were acquired.

Pre-processing of MRI data. All pre-processing was conducted using the AFNI software package (Cox, 1996) except where otherwise noted. Each EPI run for each participant was slice-time corrected and then motion corrected using the last image prior to the anatomical scan as a reference. EPI images were then coregistered to the corresponding anatomical image for each participant. Using ANTs (Avants *et al.*, 2011) nonlinear warping software, the images for each participant were warped to the Talairach brain (Talairach and Tournoux, 1988). Finally, the EPI images were converted to percent signal change normalized to the mean of each run, and then spatially smoothed using a 5 mm full-width half-maximum Gaussian kernel.

Experiment 1B

Participants. Twenty-two new participants meeting the same inclusion criteria were recruited.

Behavioral task and procedure. The behavioral task and procedure were identical to those of Experiment 1A, with the exception that participants completed 4 runs of the task, the high reward was increased to 30¢, the low reward was replaced with no reward (0¢), and responses that were incorrect or too slow followed by corresponding feedback (i.e. the words 'Incorrect' or 'Too Slow') to differentiate between errors and the absence of reward for a correct response. Therefore, the high-reward color was followed by 30¢ reward on 80% of correct trials, whereas the low-reward color was followed by 30¢ reward on 20% of correct trials (and not rewarded otherwise).

MRI data acquisition. The MRI data acquisition was identical to Experiment 1A, with the exception that four rather than five runs of the task were completed.

Pre-processing of MRI data. Preprocessing of the MRI data was identical to Experiment 1A.

Data analysis

Behavioral performance. Mean RT and accuracy were computed separately for high- and low-reward targets and compared via *t*-tests. Only correct trials were included in the RT analyses and RTs more than 2.5 SD above and below the mean for a given condition for a given participant were trimmed.

Activation of the value-driven attention network by reward. All statistical analyses on the fMRI data were performed using the AFNI software package. Data for each participant were
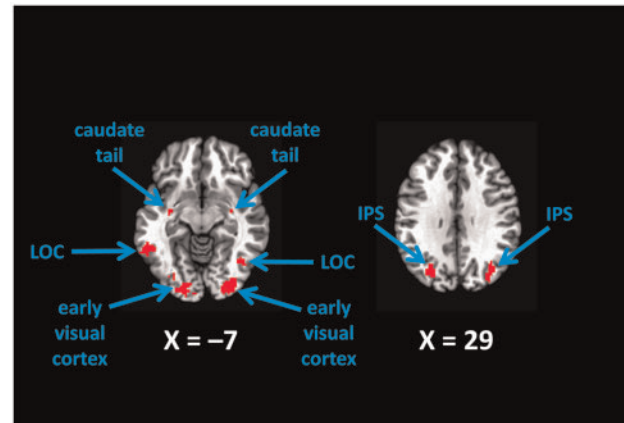

Fig. 2. The value-driven attention network as revealed by contrasting high- *vs* low-reward feedback. Voxels shown were more active in response to high *vs* low reward at P < 0.001 in each of the two experiments, reflecting a robust reward signal in visual areas. LOC, lateral occipital complex; IPS, intraparietal sulcus.

subjected to a general linear model (GLM) that included six task-related regressors: (i) high-reward target on the left, (ii) high-reward target on the right, (iii) low-reward target on the left, (iv) low-reward target on the right, (v) high-reward feedback, (vi) and low-reward feedback. These regressors were modeled using a canonical hemodynamic response function (HRF), each being represented by a simple HRF onset to the appearance of the corresponding event. Regressors of non-interest included six motion parameters and drift in the scanner signal.

The resulting beta-weight estimates for high- and low-reward feedback were compared via a t-test separately for Experiments 1A and 1B. Images were then thresholded at a voxelwise P < 0.001 uncorrected, and voxels active in *both* experiments were maintained to produce a region of interest (ROI) mask. Subsequent analyses were then performed on particular regions within this mask comprising the value-driven attention network (bilateral IPS, LOC, caudate tail, and early visual cortex) along with the caudate head and anterior cingulate cortex (ACC). These regions were identified and labeled based on their anatomical coordinates and in consideration of prior studies on the neural correlates of value-driven attention; the labels do not reflect independent functional localization and are intended for descriptive purposes (see Figure 2).

Unfolding of reward responses in the value-driven attention network over time. To examine the temporal unfolding of the observed reward responses in the value-driven attention network, the same GLM used to define these reward responses (see above) was performed separately over each run of the task. Then, the difference in the beta-weights for high- and low-reward feedback was extracted from ROIs (as defined earlier) for each run and then compared across runs via a linear trend analysis. Because Experiment 1A contained one more run than Experiment 1B, the difference scores for the last two runs in Experiment 1A were averaged for the purposes of this analysis.

Representation of cue position. To examine whether the reward signals in the value-driven attention network represent the spatial position of the preceding cue (i.e. the reward-predictive target), a different GLM was performed in which the regressors for high- and low-reward feedback were further divided based on whether the preceding target was presented on the left or right side of the display. There were now eight task-related regressors: (i—iv) the same target regressors coding value by position, (v) high-reward feedback following a target
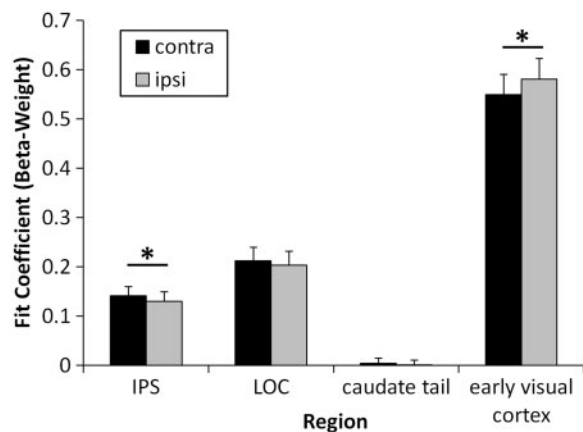
Fig. 3. Mean beta-weights for responses to high-reward feedback when the preceding target was presented in the contralateral *vs* ipsilateral hemifield by region. Reward signals in both the IPS and early visual cortex were modulated by the position of the preceding target. Error bars reflect the SEM *$P < 0.05$.

on the left, (vi) high-reward feedback following a target on the right, (vii) low-reward feedback following a target on the left and (viii) low-reward feedback following a target on the right, in addition to motion and drift regressors. Then, the beta-weights for high-reward feedback following a right- and left-lateralized target were compared within the value-driven attention network ROIs, and the same comparison was done using the beta-weights for low-reward feedback. For each region (e.g. right LOC), the difference in response to reward feedback following contralateral and ipsilateral targets was computed and then averaged across each hemisphere of a region to create a per-region contralateral *vs* ipsilateral comparison (see Figure 3).

Representation of cue identity. To examine whether the reward signals in the value-driven attention network represent the identity (in this case, color) of the preceding cue (i.e. the reward-predictive target), a different GLM was performed in which the regressors for high- and low-reward feedback were further divided based on whether the preceding target was red or green. This resulted in eight task-related regressors: (i–iv) the same target regressors coding value by position, (v) high-reward feedback following a red target, (vi) high-reward feedback following a green target, (vii) low-reward feedback following a red target and (viii) low-reward feedback following a green target, in addition to motion and drift regressors. This GLM was performed on each run separately, providing one beta-weight estimate per condition per run. Then, beta-weights for high-reward feedback following a red and green target were extracted within each (bilateral) ROI and subjected to multivariate pattern analysis (MVPA) using the linear support vector machine (SVM) functions in Matlab. Linear SVM was performed using leave-one-run-out cross-validation, such that the SVM was iteratively trained on the beta-weights from n-1 of the runs and tested on the left out run for each participant. Then, mean classification accuracy across participants was computed. This was then repeated using the beta-weights for low-reward feedback. The probability of the observed mean classification accuracy under the null hypothesis was determined using a randomization procedure in which a distribution of mean classification accuracy was computed under conditions in which the training labels were randomly shuffled for each participant (10 000 iterations).

Analysis of reward prediction errors. A similar GLM was performed as for the representation of cue identity, but in this case collapsing over all trials (rather than performed separately for

each run to facilitate MVPA), with the reward feedback coded in terms of whether it was expected or unexpected by virtue of the value assigned to the preceding target. This resulted in eight task-related regressors: (i–iv) the same target regressors coding value by position, (v) unexpected high-reward feedback, (vi) expected high-reward feedback, (vii) unexpected low-reward feedback and (viii) expected low-reward feedback, in addition to motion and drift regressors. Two subsequent contrasts were performed on the beta-weights over the four reward feedback conditions: one reflecting the interaction between value and expectation (valence-dependent prediction errors) and another reflecting the main effect of expectation (valence-independent prediction errors). Contrast images were thresholded at voxel-wise $P < 0.005$ uncorrected, and the resulting clusters were assessed for statistical significance using the AFNI program 3dClustSim (n iterations = 10 000; clusters defined using nearest neighbor method; cluster threshold: $P < 0.05$).

Results

Behavioral performance

Neither mean RT nor accuracy was significantly modulated by the value of the target, $ts < 0.60$, $Ps > 0.55$, consistent with prior studies using this paradigm (e.g. Anderson et al., 2011a, 2012, 2013). Overall performance was high (mean accuracy: 92%, mean RT: 595 ms), providing amble opportunity to experience and learn from the stimulus–reward contingencies. It should be noted that the participants in Experiment 1A completed a subsequent test phase which was the focus of a prior study demonstrating behavioral and neural evidence of attentional priority for the reward-associated colors (Anderson *et al.*, 2014), suggesting that participants do associate the targets colors with reward in this paradigm. Successful learning of the cue–reward pairings was also corroborated by an analysis of reward prediction errors (see below).

Activation of the value-driven attention network by reward

Each region of the value-driven attention network, including the caudate tail, LOC, IPS and early visual cortex, was more active bilaterally following high-reward feedback than low reward feedback (Figure 2), even though the identity and position of targets were separately modeled. These reward signals in the value-driven attention network were robust, being independently evident in each Experiment (see 'Methods' section). Reward signals were also evident outside of the value-driven attention network in regions typically associated with reward and feedback processing, including the caudate head (e.g. Kim and Hikosaka, 2013; Hikosaka et al., 2014) and ACC (e.g. Carter *et al.*, 1998; Kiehl *et al.*, 2000; Hickey *et al.*, 2010), which were not of interest in this study and were not further analyzed. The reward signals in the value-driven attention network gradually increased over the course of the experiment, exhibiting a linear trend over runs of the task, $F(1,39) = 5.42$, $P = 0.025$, $\eta_p^2 = 0.057$, consistent with a learning-dependent representation. By comparison, such a trend was not observed in the caudate head or ACC, $Fs < 0.25$, $Ps > 0.60$, consistent with their role in the online monitoring of feedback (e.g. Carter *et al.*, 1998; Kiehl *et al.*, 2000; Hickey *et al.*, 2010) and in representing current rewards (Kim and Hikosaka, 2013; Hikosaka et al., 2014).

Stimulus specificity of the reward signals

Spatial position. A 4 × 2 analysis of variance (ANOVA) with region (caudate tail, LOC, IPS and early visual cortex) and preceding target location (contralateral vs ipsilateral) as factors was performed on the beta-weights for high-reward feedback (see 'Methods' section). This analysis revealed a main effect of region, $F(3,117) = 99.30$, $P < 0.001$, $\eta_p^2 = 0.718$, reflecting stronger reward responses in early visual cortex. There was no main effect of location, $F(1,39) = 0.14$, $P = 0.713$; however, location interacted robustly with region, $F(3,117) = 6.48$, $P < 0.001$, $\eta_p^2 = 0.142$ (Figure 3). Pairwise comparisons revealed greater activation by reward when the preceding target was presented in the contralateral hemifield for the IPS, $t(39) = 2.55$, $P = 0.015$, $d = 0.40$, whereas early visual cortex showed the opposite pattern, $t(39) = -3.19$, $P = 0.003$, $d = 0.50$. LOC and the caudate tail showed no evidence of spatial encoding of the prior target, $ts < 0.84$, $Ps > 0.40$. The same analysis performed on the mean beta-weights for low-reward feedback revealed no reliable interaction between region and location, $F(3,117) = 1.30$, $P = 0.279$, although there was a trend towards the same effect in early visual cortex, $t(39) = -1.91$, $P = 0.064$ (other regions: $ts < 0.68$, $Ps > 0.30$).

Object identity. For the activity evoked by high-reward feedback, the color of the preceding target could be successfully decoded in LOC (58.2%, $P < 0.001$), the caudate tail (56.3%, $P = 0.004$), and early visual cortex (56.8%, $P = 0.001$), but not IPS (52.1%, $P = 0.179$). Classification in the LOC, caudate tail, and early visual cortex remain significant when applying Bonferroni correction for multiple comparisons ($\alpha = 0.0125$). For the activity evoked by low-reward feedback, decoding of the preceding target color was unsuccessful; although there was a trend towards successful classification in early visual cortex (54.1%, $P = 0.03$), this did not pass correction for multiple comparisons (other regions < 52.7%, $Ps > 0.12$ uncorrected).

Analysis of reward prediction errors

To confirm learning of the stimulus–reward associations, beta-weights for reward feedback were computed in terms of value (i.e. the magnitude of the reward) and expectation (i.e. whether a reward of that magnitude would have been expected or unexpected given the probabilities assigned to the preceding target). If participants had learned the stimulus–reward contingencies, then reward prediction errors should be evident. An interaction between value and expectation could be seen in right orbitofrontal cortex (OFC) (see Figure 4), consistent with the coding of reward prediction errors in this region (e.g. O'Doherty, 2004); that is, the difference in the response evoked by high vs low reward was greater when these values were unexpected than when they were expected given the preceding target. A main effect of expectation was also evident in a region encompassing the posterior cingulate cortex (PCC) and ventral precuneus, with expected reward eliciting greater responding than unexpected reward (see Pearson et al., 2011).

Discussion

The signaling of value-based attentional priority is reflected in the activation of a network of brain regions including the LOC, IPS, caudate tail and early visual cortex by a learned reward cue (see Anderson, 2016, for a review). Although this value-driven attention network is known to play a role in attentional orienting once stimulus–reward relationships have been learned, how
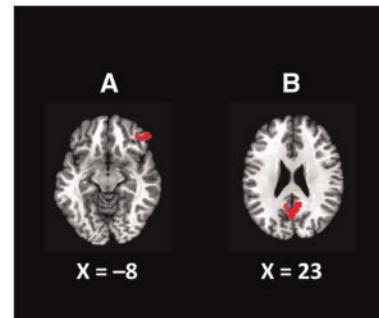

Fig. 4. Regions in which reward prediction error signals were evident. **(A)** A significant interaction between value (i.e. the magnitude of reward feedback) and expectation (based on the reward probabilities assigned to the target color) was evident in the right OFC. **(B)** A main effect of expectation, with a stronger response for expected rewards, was evident in a region encompassing the posterior cingulate and ventral precuneus.

signals combining visual information with value information arise in these brain areas during the course of learning remains unaddressed. Reward prediction-like signals have been observed in early visual cortex (Shuler and Bear, 2006; Arsenault et al., 2013), potentially reflecting a broader principle that could mediate reward-induced plasticity within the visual system (Roelfsema and van Ooyen, 2005; Rombouts et al., 2015), and reward-prediction signals are known to play an important role in the development of attentional biases towards reward cues (Anderson et al., 2013; Sali et al., 2014). However, the stimulus specificity of reward signals in visual areas is not known, nor has the extent of such signals throughout a broader attention network been examined.

Integration of cue and value information

In this study, I observed robust reward signals throughout the value-driven attention network, extending prior findings in early visual cortex (Shuler and Bear, 2006; Arsenault et al., 2013). These reward signals became stronger with experience with the stimulus–reward associations, the learning of which was confirmed via reward prediction error signals in other regions of the brain. Importantly, these reward signals within the value-driven attention network carried information about the features of the preceding reward cue, namely its position and identity. The observed reward signals therefore represent the key components necessary for signaling value-based attentional priority (e.g. Anderson et al., 2011b, 2014; Anderson, 2013), reflecting a candidate reinforcement signal that could serve to potentiate stimulus-specific representations as a function of their ability to predict reward.

My findings also suggest functional specialization within the value-driven attention network during the process of learning stimulus–reward associations. Specifically, although both the position and identity of reward cues is represented in each of the regions of the value-driven attention network during the capture of attention after reward learning has occurred (Peck et al., 2009; Qi et al., 2013; Yamamoto et al., 2013; Anderson et al., 2014; MacLean and Giesbrecht, 2015), such information is separately represented along the dorsal/ventral stream during learning. Specifically, the IPS represents spatial position but not object identity, whereas the opposite is true of the LOC and caudate tail, reflecting known principles of functional specialization within the visual system (e.g. Gill-Spector et al., 2001; Corbetta and Shulman, 2002; Bisley and Goldberg, 2003; Balan

and Gottlieb, 2006; Yamamoto *et al.*, 2012); both receive input from early visual cortex (e.g. Corbetta and Shulman, 2002; Seger, 2013), which was found to represent both position and identity during learning. Value-based attentional priority reflects the integration of these reinforcement signals throughout the value-driven attention network, creating a bound representation that is sufficient to elicit stimulus-specific orienting.

Limitations and future directions

The existence of value-based attentional priority signals assumes some integration of value information with visual information during learning in the areas involved in signaling such priority, but the mechanisms of such integration have remained unaddressed. Although the present study sheds new light on this issue by outlining a clear example of such integration, the findings should not alone be taken as an explanation for the development of value-dependent attentional bias. Future research is needed to iron out the dynamics of these candidate teaching signals and how they might be translated into an enduring bias.

The degree to which the observed reward signals depend on task-specific motivation is also unclear. The reward cues are task-relevant targets in this study and in the training phase of many other studies examining attentional biases for learned reward cues (see Anderson, 2013, 2016, for reviews). However, value-dependent attentional biases have been observed for stimuli that were never presented as targets but none-the-less predict reward (e.g. Le Pelley *et al.*, 2015). Thus, the observed reward signals could reflect voluntary and motivated learning, or automatic associative learning. The relationship between voluntary processes and the development of automatic processes remains a largely open question in research on value-driven attention. Similarly, these signals could depend on covert selection of the prior target, overt selection, both, or neither (i.e. independent of prior selection), as fixation was not enforced in this experiment. However, spatial content was clearly evident.

The ability to decode previously experienced cue features during feedback processing was evident predominantly following high reward, suggesting a reward-mediated process rather than a more general principle of feedback processing or a consequence of prior target selection per se (such that the nature or presence of feedback is superfluous). Given the probabilistic reward structure with two levels of reward, complete learning of the relative value of the targets is possible when anchoring learning specifically to the probability of high-reward feedback. Examination of the nature of stimulus-specific feedback signals in the context of more complex reward structures (e.g. Navalpakkam *et al.*, 2010) would be interesting and could reveal a richer source of value-dependence.

Interestingly, the spatially-specific reward signals in early visual cortex reflected preferential responding in the hemisphere ipsilateral to where the reward cue was presented, which contrasts with the contralateral bias observed in IPS and typically thought to reflect an enhancement of stimulus-evoked activity (e.g. Anderson *et al.*, 2014; Peck *et al.*, 2009; Qi *et al.*, 2013). Although somewhat paradoxical, the reward responses in early visual cortex clearly contained spatial information about the preceding cue and replicate the pattern previously observed by Arsenault *et al.* (2013). In that study, unexpected reward generated a prediction error signal that reflected suppression of voxels sensitive to the reward cue in early visual cortex. Reduced activation in contlateral compared with ipsilateral visual cortex could also be explained by suppression of the reward cue representation, which might reflect a broader principle governing plasticity within this brain region. Perhaps such suppression serves in the interest of reducing interference from on-going visual input during the process of consolidating visual memory at later stages of processing (including in LOC, IPS and caudate tail). The significance of such seemingly paradoxical reward signals in the early visual cortex is an intriguing question that clearly warrants further investigation.

Conclusions

The findings from the present study highlight a mechanism by which value information is combined with visual information in the human attention system. When high rewards are received, the feature representations activated by the preceding cue become re-activated (or suppressed). These reward signals are evident throughout the visual system, within a network of brain regions corresponding very closely to those known to be involved in attentional orienting to learned reward cues (e.g. Peck *et al.*, 2009; Yamamoto *et al.*, 2013; Anderson *et al.*, 2014 Anderson, 2016), which could drive the plasticity hypothesized to underlie value-driven attention (Rombouts *et al.*, 2015; Anderson, 2016). Such reward signals could also conceivably have an immediate impact on stimulus priming, thereby providing a common mechanism linking value-driven attentional capture by learned reward cues (Anderson *et al.*, 2011b, 2014; Anderson, 2013, 2016) and reward-mediated priming (Hickey *et al.*, 2010, 2014; Hickey and van Zoest, 2012, 2013).

References

Anderson, B.A. (2013). A value-driven mechanism of attentional selection. *Journal of Vision*, **13**(3), 7.

Anderson, B.A. (2016). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences*, **1369**, 24–39.

Anderson, B.A., Kuwabara, H., Wong, D.F., *et al.* (2016). The role of dopamine in value-based attentional orienting. *Current Biology*, **26**, 550–5.

Anderson, B.A., Laurent, P.A., Yantis, S. (2011a). Learned value magnifies salience-based attentional capture. *PLoS One*, **6**(11),e27926.

Anderson, B.A., Laurent, P.A., Yantis, S. (2011b). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 10367–71.

Anderson, B.A., Laurent, P.A., Yantis, S. (2012). Generalization of value-based attentional priority. *Visual Cognition*, **20**, 647–58.

Anderson, B.A., Laurent, P.A., Yantis, S. (2013). Reward predictions bias attentional selection. *Frontiers in Human Neuroscience*, **7**, 262.

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.

Arsenault, J.T., Nelissen, K., Jarraya, B., Vanduffel, W. (2013). Dopaminergic reward signals selectively decrease fMRI activity in primate visual cortex. *Neuron*, **77**, 1174–86.

Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, **54**, 2033–44.

Balan, P.F., Gottlieb, J. (2006). Integration of exogenous input into a dynamic salience map revealed by perturbing attention. *Journal of Neuroscience*, **26**, 9239–49.

Bisley, J.W., Goldberg, M.E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science*, **299**, 81–6.

Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433–6.

Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., Cohen, J.D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, **280**, 747–9.

Corbetta, M., Patel, G., Shulman, G.L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, **58**, 306–24.

Corbetta, M., Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, **3**, 201–15.

Cox, R.W. (1996). AFNI: software for the analysis and visualization of function magnetic resonance neuroimages. *Computers and Biomedical Research*, **29**, 162–73.

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**, 529–41.

Gill-Spector, K., Kourtzi, Z., Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, **41**, 1409–22.

Hickey, C., Chelazzi, L., Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience*, **30**, 11096–103.

Hickey, C., Chelazzi, L., Theeuwes, J. (2014). Reward priming of location in visual search. *PLOS One*, **9**, e103372.

Hickey, C., Peelen, M.V. (2015). Neural mechanisms of incentive salience in naturalistic human vision. *Neuron*, **85**, 512–8.

Hickey, C., van Zoest, W. (2012). Reward creates oculomotor salience. *Current Biology*, **22**, R219–20.

Hickey, C., van Zoest, W. (2013). Reward-associated stimuli capture the eyes in spite of strategic attentional set. *Vision Research*, **92**, 67–74.

Hikosaka, O., Kim, H.F., Yasuda, M., Yamamoto, S. (2014). Basal ganglia circuits for reward value guided behavior. *Annual Review of Neuroscience*, **37**, 289–306.

Hikosaka, O., Yamamoto, S., Yasuda, M., Kim, H.F. (2013). Why skill matters. *Trends in Cognitive Sciences*, **17**, 434–41.

Hopf, J.M., Schoenfeld, M.A., Buschschulte, A., Rautzenberg, A., Krebs, R.M., Boehler, C.N. (2015). The modulatory impact of reward and attention on global feature selection in human visual cortex. *Visual Cognition*, **23**, 229–48.

Kiehl, K.A., Liddle, P.F., Hopfinger, J.B. (2000). Error processing in the rostral anterior cingulate: an event-related fMRI study. *Psychophysiology*, **37**, 216–23.

Kim, H.F., Hikosaka, O. (2013). Distinct basal ganglia circuits controlling behaviors guided by flexible and stable values. *Neuron*, **79**, 1001–10.

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**, 158–71.

MacLean, M.H., Giesbrecht, B. (2015). Neural evidence reveals the rapid effects of reward history on selective attention. *Brain Research*, **1606**, 86–94.

Navalpakkam, V., Koch, C., Rangel, A., Perona, P. (2010). Optimal reward harvesting in complex perceptual environments. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 5232–7.

O'Doherty, J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology*, **14**, 769–76.

Pearson, J.M., Heilbronner, S.R., Barack, D.L., Hayden, B.Y., Platt, M.L. (2011). Posterior cingulate cortex: adapting behavior to a changing world. *Trends in Cognitive Sciences*, **15**, 143–51.

Peck, C.J., Jangraw, D.C., Suzuki, M., Efem, R., Gottlieb, J. (2009). Reward modulates attention independently of action value in posterior parietal cortex. *Journal of Neuroscience*, **29**, 11182–91.

Qi, S., Zeng, Q., Ding, C., Li, H. (2013). Neural correlates of reward-driven attentional capture in visual search. *Brain Research*, **1532**, 32–43.

Roelfsema, P.R., van Ooyen, A. (2005). Attention-gated reinforcement learning of internal representations for classification. *Neural Computation*, **17**, 2176–214.

Rombouts, J.O., Bohte, S.M., Martinez-Trujillo, J., Roelfsema, P.R. (2015). A learning rule that explains how rewards teach attention. *Visual Cognition*, **23**, 179–205.

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**, 1654–64.

Seitz, A.R., Kim, D., Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, **61**, 700–7.

Seger, C.A. (2013). The visual corticostriatal loop through the tail of the caudate: circuitry and function. *Frontiers in Systems Neuroscience*, **7**, 104.

Shuler, M.G., Bear, M.F. (2006). Reward timing in the primary visual cortex. *Science*, **311**, 1606–9.

Talairach, J., Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme.

van Koningsbruggen, M.G., Ficarella, S.C., Battelli, L., Hickey, C. (2016). Transcranial random noise stimulation of visual cortex potentiates value-driven attentional capture. *Social, Cognitive, and Affective Neuroscience*, **11**, 1481–8.

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**, 11227–38.

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*, **32**, 11005–16.