the british
psychological society
promoting excellence in psychology

British Journal of Psychology (2021)

www.wileyonlinelibrary.com

Reward produces learning of a consciously inaccessible feature

Xue Dong[1,2], Mingxia Zhang[1,2], Bo Dong[1,3], Yi Jiang[2,4,5,6,7] and Min Bao[1,2]*

[1]CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
[2]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China
[3]Department of Psychology, Suzhou University of Science and Technology, Suzhou, China
[4]State Key Laboratory of Brain and Cognitive Science, Beijing, China
[5]CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai, China
[6]Chinese Institute for Brain Research, Beijing, China
[7]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, China

Reward has a significant impact on behaviour and perception. Most past work in associative reward learning has used perceptually distinct visual cues to associate with different reward values. Thus, it remains unknown to what extent the learned bias towards reward-associated stimuli depends on consciousness of the apparent differences between stimuli. Here, we resolved this issue by using an inter-ocular suppression paradigm with the monetary rewarding and non-rewarding cues identical to each other except for their eye-of-origin information. Thus, the reward coding system cannot rely on consciousness to select the reward-associated cue. Surprisingly, the targets in the rewarded eye broke into awareness faster than those in the non-rewarded eye. We further revealed that producing this effect required both top-down attention and inter-ocular suppression. These findings suggest that the human's reward coding system can produce two different types of reward-based learning. One is independent of consciousness yet fairly consuming attentional resources. The other one results from volitional selection of stimuli of behavioural significance.

Actions or perceptions can be biased in favour of reward-associated stimuli (Anderson, Laurent, & Yantis, 2011b; Hickey, Chelazzi, & Theeuwes, 2010; Proshansky & Murphy, 1942; Thorndike, 1911). For example, task-irrelevant distractors previously associated with high reward slow visual search more than the equally salient distractors previously associated with low reward (Anderson, Laurent, & Yantis, 2011a). Furthermore, facial stimuli associated with learned reward value can survive the attentional blink (Raymond & O'Brien, 2009). Later studies reported that rewarded percepts more frequently dominate

---

*Correspondence should be addressed to Min Bao, CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Road, Chaoyang District, Beijing 100101, China (email: baom@psych.ac.cn).

DOI:10.1111/bjop.12518

awareness than non-rewarded or punished percept in a binocular rivalry task (Marx & Einhauser, 2015; Wilbertz, van Slooten, & Sterzer, 2014). All these findings indicate the privileged processing of reward-associated stimuli.

It is worth noting that in most of the previous studies, the rewarding vs. non-rewarding (or high vs. low rewarding) stimuli are consciously perceived and have distinguishable appearances (Marx & Einhauser, 2015; Raymond & O'Brien, 2009; Wilbertz et al., 2014; Xue, Zhou, & Li, 2015). Therefore, it seems plausible that the biased response following reward learning depends on consciousness of the apparent differences between the stimuli. Nevertheless, it is already known that neuromodulatory signals for rewards (and punishments) are released diffusely throughout the entire brain (Dalley et al., 2001; Schultz, 2000; Vickery, Chun, & Lee, 2011). Intuitively, the effects of reward learning could be free from the constraints of consciousness. Unfortunately, to our knowledge, this hypothesis cannot be easily proved by the previous work, because in most of them perceptually distinguishable visual cues are used to associate with different reward values.

The present study introduces a novel paradigm in which the participants cannot consciously differentiate the monetary rewarding and non-rewarding visual cues. This was realized by rendering the two visual cues identical to each other except for their eye-of-origin information. Specifically, we adopted a b-CFS paradigm (Jiang, Costello, & He, 2010). By this paradigm, one can present a target in one eye and a dynamic sequence of complex, geometric images (i.e., continuous flash suppression stimuli, CFS) in the other eye. Usually, the target is rendered invisible due to the suppression from the CFS stimuli, a visual phenomenon called inter-ocular suppression. However, given longer presentation, the suppression will become ineffective. Therefore, the target may break into awareness, at which time the participant is required to immediately press a key to report seeing the target (that is why the paradigm is often called b-CFS). In the present study, participants were rewarded only when they reported seeing the target in one of the two eyes, which we called the rewarded eye. Because the targets presented to both rewarded eye and non-rewarded eye were of the same appearance except for their eye-of-origin information, participants should have no conscious knowledge about the difference between the targets in the two eyes (Wolfe & Franzel, 1988; Zhang, Jiang, & He, 2012). Thus, any eye-specific learning effects established over the reward-based training should be contributed from the reward coding system, but independent of the consciously perceived difference of the two kinds of targets. Experiment 1 proved that the eye-specific reward learning could be observed using the b-CFS paradigm. Experiment 2 demonstrated that inter-ocular suppression was necessary for the learning effect to occur. Experiment 3 further revealed that when reward was related to the combination of a consciously inaccessible feature (i.e., eye of origin) and another consciously accessible feature of target, learning mainly relied on the consciously accessible cue. Finally, Experiment 4 examined whether participants could realize the difference between rewarding and non-rewarding targets. Note that in this article, to avoid confusion, the term 'consciousness' is specifically used to describe whether participants can discern the difference between the rewarding and non-rewarding targets, whereas 'awareness' is used only in the situation to describe that stimulus in one eye is dominating the current perception when stimulus in the other eye is temporarily rendered invisible.

Experiment 1: Eye-specific reward learning

We first examined our original hypothesis about the reward-induced eye-specific learning to consciously inaccessible stimuli and the role of top-down attention in Experiments 1a and 1b, respectively.

Method

Participants

Participants performed a screen test in experiments 1-3 before the formal experiments. The goal was to find 1) individuals with relatively balanced ocular dominance and 2) the stimulus contrast for non-extreme breakthrough ratio for each eye (see the procedures for detail), since any ceiling or floor effects on the measurements of breakthrough ratio could hamper the observation of any learning effect.

Thirty-six participants (17 males and 19 females) finished Experiment 1a. They were screened from 105 volunteers. We initially tested the reward learning effect on 9 participants and used G*power (Faul, Erdfelder, Buchner, & Lang, 2009) to estimate the required sample size for statistical power of 0.8. The minimum required sample size was 13. Thus, to confirm the novel finding in Experiment 1a, another two experimenters, who collected data from 13 and 14 participants, respectively, replicated the experiment. The later results were similar with our initial finding; therefore, we combined data of all 36 participants together in the analysis. Besides, there were at least 14 participants in each of the following experiments. Fourteen participants (7 males and 7 females) who were screened from 45 volunteers finished Experiment 1b. The participants ranged in age from 18 to 27 years, and all had normal or corrected-to-normal vision and were naïve to the experimental hypotheses. The experiments were conducted in accordance with the Declaration of Helsinki. Experimental procedures were approved by the Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences, and informed consent was obtained from each participant.

Apparatus

Stimuli were presented on a 21-in Dell CRT monitor with a resolution of 1024 × 768 pixels at a refresh rate of 85 Hz and programmed in MATLAB and Psychtoolbox-3 (Brainard, 1997). The display was calibrated with a Photo Research PR-655 spectrophotometer. To calibrate the display, we measured the luminance gamma curves and inverted them with a look-up table. The mean luminance of the screen was 50.9 cd/m^2. A chin-rest was used to help minimize head movement.

Stimuli and procedures

Experiment 1a. Stimuli were presented on a mid-grey background. The target was a dark grey square frame (1.2° × 1.2°, linewidth: 0.11°) with a horizontal or vertical bar (length: 0.8°, linewidth: 0.11°) in the centre. The target was displayed foveally in one eye, centred 0.25° away from the central fixation point (0.2°). The central bar could be vertically oriented to the left or right of the fixation or horizontally oriented above or below the fixation (see Figure 1A-B). The CFS stimuli (8° × 8°, flashing at 10 Hz) were displayed foveally in the other eye, which consisted of 60 images created by drawing rectangles of random colours and sizes. The black-and-white square frame (11° × 11°, linewidth: 0.11°) and the central fixation were always presented to two eyes simultaneously to help fusion.

Each trial started with a presentation of the central fixation point for 800 ms. Afterwards, the CFS stimuli appeared in one eye and kept flashing until the end of the trial. The target appeared in the other eye after a random interval varied between 100 and 400 ms. The contrast of the target ramped up to its highest level within the initial 1500 ms and then remained at the highest contrast for 500 ms. Participants were required to report the position of the central bar of the target relative to the fixation once the targets broke into awareness. They were told to respond as quickly as possible on the premise of accuracy. The trial terminated once a response was made, otherwise the target would be displayed for 2000 ms in total followed by a 600-ms blank interval while the CFS stimuli were still presented in the other eye. To prevent any afterimage of the target, the target kept drifting back and forth for 0.08° at 1 Hz along the diagonal.

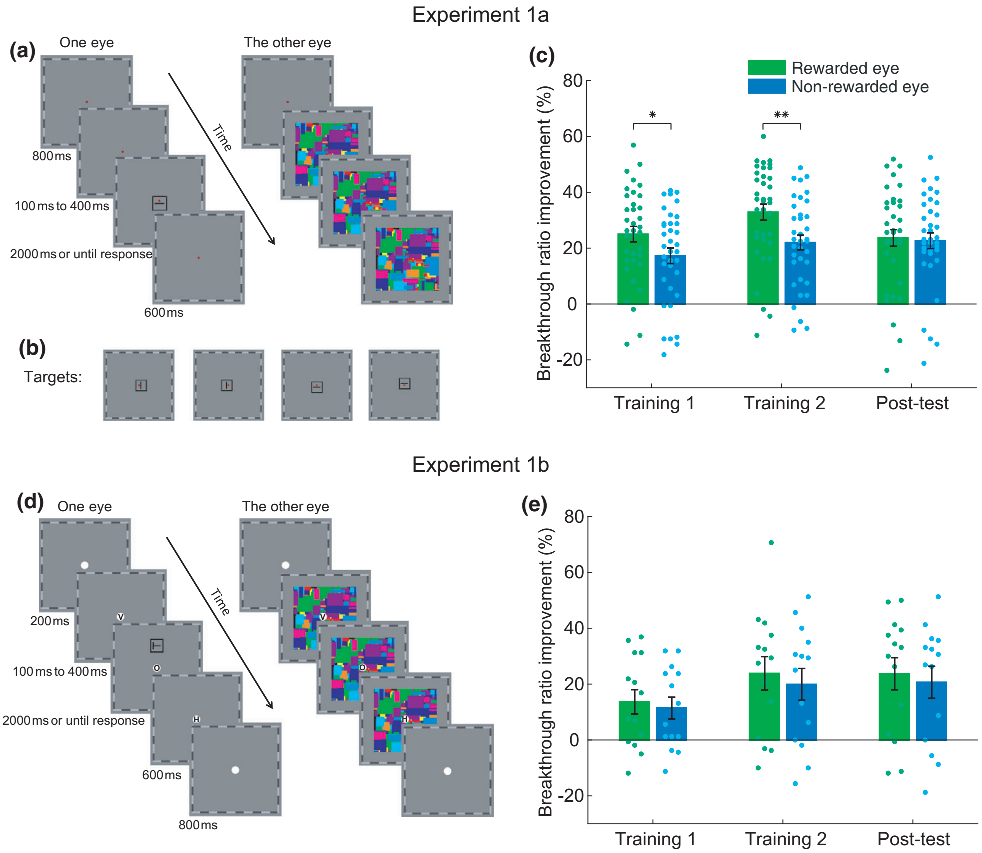

Figure 1. Stimuli and results of Experiment 1. (A) Trial sequence. (B) The examples of all possible presentations of targets in Experiment 1a, corresponding to the keypress of RightArrow, LeftArrow, DownArrow, and UpArrow (C) The improvement of breakthrough ratio relative to pre-test in training and post-test sessions in Experiment 1a. (D) The stimuli and trial sequence in Experiment 1b. (E) The improvement of breakthrough ratio relative to pre-test in training and post-test sessions in Experiment 1b. Error bars show ±1 *SEM*. Asterisks indicate the significance level (with $^{*}p_{\mathrm{FDRcorrected}} < .05$ and $^{**}p_{\mathrm{FDRcorrected}} < .01$).

Each volunteer first participated in a screen test. We used a criterion that was decided in a preliminary stage of the study. Specifically, for targets of a certain contrast, the difference of breakthrough ratio across the eyes should not exceed 20%, while the breakthrough ratio for each eye had to be between 20% and 60%. With this criterion (around 40%), the estimation of training-induced increment of breakthrough ratio may not be easily affected by a floor effect or a ceiling effect from repeatedly performing a b-CFS task. Eventually, the breakthrough ratios of 29/36 of the participants were within this range. For the other 7 participants, we used a less stringent criterion (18% - 65%) to have a larger sample. The contrasts for the target and CFS determined by the screen test were then used in the subsequent formal experimental sessions. The optimal contrasts used for each participant are listed in Table S2, as well as the breakthrough ratios for each eye.

After the pre-test, the participants completed two training sessions and a post-test. Each session consisted of two blocks of 160 trials. All the tasks were completed in one day with short breaks between sessions. In the pre- and post-tests, there were no monetary rewards. The target was presented to the left eye in half of the trials and to the right eye in the rest of the trials. The two conditions of trials were randomly interleaved. In the training sessions, for each participant one eye was assigned to be the rewarded eye. Participants were not aware of this setting. The selection of the rewarded eye was counter-balanced across participants. A trial was called a rewarding trial if the target was presented to the rewarded eye. Immediately after a correct response for a rewarding trial, a 500-Hz tone would beep for 50 ms, which notified the participants of winning 0.2 yuan. After each block, a message on the screen showed participants the total amount of gain. For each session, the breakthrough ratio was calculated by dividing the number of trials with correct responses by the total trial count for each eye-of-origin condition, respectively.

Experiment 1b. The stimuli of the b-CFS task were similar to those in Experiment 1a, except that the target was a capital letter ‘T’ in a squared frame. The target was presented at 2° eccentricity above or below the centre of the screen (Figure 1D). The letter had four orientations (upright, upside down, right tilt, and left tilt). Participants were asked to report the letter orientation by pressing the corresponding arrow key. Simultaneously with the b-CFS task, participants were required to complete a central RSVP task. A series of capital letters were presented in a central white circle (0.6° in diameter) to both eyes during the presentation of CFS stimuli. Each letter subtended for 0.5° and was presented for 250 ms. The task was to press SPACE after finding ‘O’ in the letter series. A trial lasted for 3700–4000 ms or until the press of an arrow key was detected. Only one letter ‘O’ was presented in each trial, and it was not presented in the first or the last 200 ms. In each block, 20 trials were planned to be catch trials without RSVP target. However, since a trial may end before the presentation of ‘O’, there would be more catch trials. We calculated the hit rate and false alarm rate to measure the performance of RSVP task. A hit was the response that was made after the presentation of ‘O’ and before the end of the trial. A false alarm was the response that was made before or without the presentation of ‘O’ in a trial.

The screen test and procedure were same as those in Experiment 1a. In the training sessions, participants were told that the reward they could receive firstly relied on the b-CFS task, while the hit rate of the RSVP task would serve as a discount ratio to the overall reward.

Results and discussion

Repeated measurements ANOVA and paired *t*-tests were used to statistically analyse the results (see Table S1 for the detailed statistics of the ANOVA results). Consistent with the previous finding (Mastropasqua, Tse, & Turatto, 2015), the targets broke into awareness generally faster in the later sessions than in the pre-test, as revealed by the significant main effect of session (pre-test vs. post-test, $F(1,35) = 37.80$, $p < .001$, $\eta^2 = .52$). However, neither the main effect of eye (rewarded vs. non-rewarded, $F(1,35) = .46$, $p = .503$, $\eta^2 = .01$) nor the interaction ($F(1,35) = .47$, $p = .496$, $\eta^2 = .01$) was significant. Thus, there was not any significant difference in the breakthrough time between the two eyes in either pre- or post-test.

We then analysed the results of breakthrough ratios. In the pre-test, there was no significant difference between the breakthrough ratios for the rewarded eye and the non-rewarded eye ($t(35) = .19$, $p = .848$, Cohen's $d = .03$). Therefore, we subtracted the breakthrough ratios in the pre-test from those in the subsequent sessions to estimate the change of breakthrough ratios across sessions. Surprisingly, there were more breakthrough trials in the rewarded eye than in the non-rewarded eye in the training sessions (training 1: $t(35) = 2.82$, $p_{\text{uncorrected}} = .008$, $p_{\text{FDRcorrected}} = .012$, $d = .46$ and training 2: $t(35) = 3.65$, $p_{\text{uncorrected}} < .001$, $p_{\text{FDRcorrected}} = .003$, $d = .66$, see Figure 1C). However, this eye-specific effect was absent in the post-test ($t(35) = .44$, $p_{\text{uncorrected}} = .664$, $p_{\text{FDRcorrected}} = .664$, $d = .06$) where the participants no longer received monetary rewards. In addition, we found that this eye-specific effect developed very fast. It could be observed in the first block of training (rewarded eye vs. non-rewarded eye: $t(35) = 2.28$, $p = .029$, $d = .35$).

We further examined any potential influences of false alarms. In the present study, a false alarm represents that the participants made a wrong keypress in a trial. This could be due to a wrong judgement of the target position after the targets broke into awareness or a cheating response by randomly guessing the location of the invisible target. In the training sessions, the participants might have the motivation to cheat in order to receive more monetary rewards. Considering the four-alternative forced-choice (4AFC) task, the probability of correctly guessed trials was 1/3 of that of wrongly guessed trials. Assuming that all the wrong responses were due to failed guesses, the maximum number of correctly guessed trials in theory could be estimated by dividing the number of trials with wrong responses by three. Therefore, to correct for the effects of false alarms, the maximum number of correctly guessed trials was subtracted from the number of trials with correct responses before calculating the breakthrough ratios. After the correction, the breakthrough ratio also showed more increase in the rewarded eye than in the non-rewarded eye (training 1: $t(35) = 2.78$, $p_{\text{uncorrected}} = .009$, $p_{\text{FDRcorrected}} = .013$, $d = .47$ and training 2: $t(35) = 3.56$, $p_{\text{uncorrected}} = .001$, $p_{\text{FDRcorrected}} = .003$, $d = .65$). The eye-specific effect was absent in the post-test ($t(35) = .33$, $p_{\text{uncorrected}} = .744$, $p_{\text{FDRcorrected}} = .744$, $d = .04$).

To test whether top-down attention could affect the learning process, another group of participants were asked to perform a RSVP task simultaneously with the b-CFS task in Experiment 1b. By using the central RSVP task and moving the b-CFS targets away from the central fixation position, attention deployed on the b-CFS targets should be greatly reduced as compared to our initial experiment. Participants performed well in the RSVP task in all the sessions (hit rate: 87.23 ± 8.33%, false alarm rate: 12.44 ± 19.62%). The results of the b-CFS task (Figure 1E), however, showed no significant difference between the two eyes in the breakthrough ratio of pre-test ($t(13) = .64$, $p = .536$, $d = .18$) or the improvements of later sessions (training 1: $t(13) = .76$, $p_{\text{uncorrected}} = .458$,

$p_{\text{FDRcorrected}} = .458$, $d = .15$; training 2: $t(13) = 1.14$, $p_{\text{uncorrected}} = .274$, $p_{\text{FDRcorrected}} = .412$, $d = .18$; and post-test: $t(13) = 1.62$, $p_{\text{uncorrected}} = .130$, $p_{\text{FDRcorrected}} = .391$, $d = .14$). Similar results were found after the false alarm correction.

Experiment 2: Monocular reward learning without inter-ocular suppression

As we expected, in Experiment 1a reward facilitated the detection of target presented to the rewarded eye more than that to the non-rewarded eye. However, the findings were based on the b-CFS paradigm and inter-ocular suppression. We then examined in Experiment 2a-2c whether such a paradigm was necessary or not for observing the eye-specific learning effect. Experiment 2a tested whether reward could facilitate the detection of targets in the rewarded eye when there were no CFS stimuli in the other eye. Experiment 2b and 2c examined whether sensitivity to a grating patch presented in the rewarded eye could be better than that to the non-rewarded eye.

Method

Participants

Fifteen participants (5 males and 10 females) who were screened from 57 volunteers finished Experiment 2a. Two groups of fourteen participants finished Experiment 2b (6 males and 8 females) and Experiment 2c (4 males and 10 females). The participants ranged in age from 18 to 28 years.

Stimuli and procedures

Experiment 2a. The stimuli and task were similar to Experiment 1a. However, there were two different types of trials in the training sessions, with-CFS trial and target-only trial. The with-CFS trial was identical to a typical trial in Experiment 1a. In a target-only trial, no CFS stimuli were presented, and a target ramped up from 0 to −0.8 contrast (Weber contrast, $C = (L_s - L_b) / L_b$, where L_s and L_b denoted the luminance of the stimulus and background) within 2000 ms in one of the two eyes.

The screen test and pre-test were the same as those in Experiment 1a. During the training sessions, one eye was assigned to be the rewarded eye. However, rewards just occurred in the *target-only* trials where the target appeared in the rewarded eye. A correct response in a rewarding trial would bring a 500-Hz beep and give rise to a reward of 0.31 yuan. In this experiment, all the participants could finish the task with nearly perfect performance. Therefore, in order to make the total amount of rewards slightly different across the participants, a random amount (ranging from −5.00 to 5.00 yuan) was added to the final gain for each block before it was shown on the screen. Each participant completed a pre-test and three training sessions.

Experiment 2b. The stimuli were sinusoidal gratings (3° in diameter, 1.5 cpd), presented on the centre of a mid-grey background for only one eye (monocular) (Figure 2B). The orientation of gratings was fixed for each participant (either vertical or horizontal), but was counter-balanced across the participants.

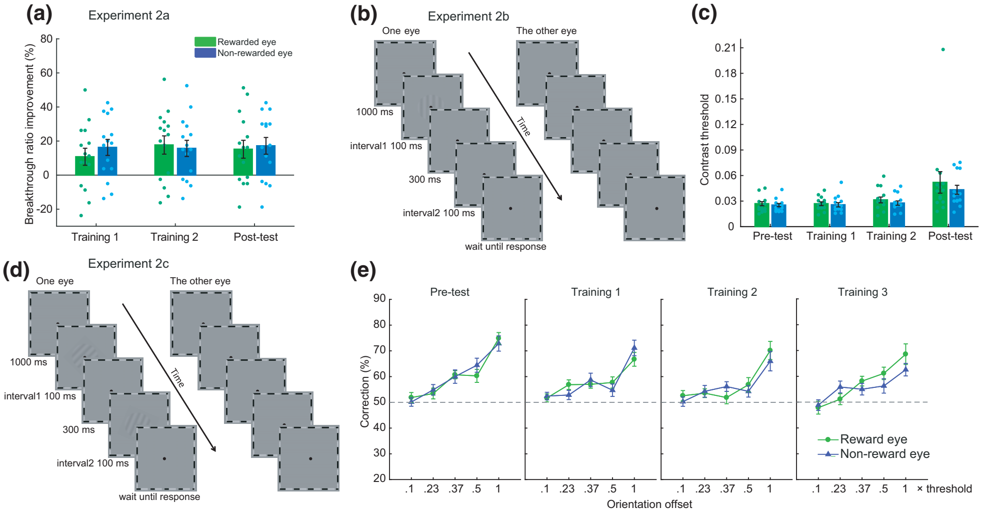

Figure 2. Stimuli and results of Experiment 2. (A) The improvement of breakthrough ratio relative to pre-test in three training sessions of Experiment 2a. (B) Trial sequence of the contrast detection task in Experiment 2b. (C) Contrast detection thresholds of four task sessions of Experiment 2b. (D) Trial sequence of the orientation discrimination task in Experiment 2c. (E) Orientation discrimination thresholds of four task sessions of Experiment 2c.

Participants performed a two-interval forced-choice (2IFC) task. Every trial started with a 1000-ms blank, followed by two 100-ms test intervals which were separated by a 300-ms gap. Each interval was signalled by a tone. Participants were required to detect in which interval the grating was presented. Three practice sessions and four formal test sessions were completed. The contrasts of test gratings were manipulated by 2-down-1-up staircases in the practice sessions. Sixty contrast levels were predetermined for the staircase, ranging logarithmically from 0.4% (Michelson contrast, $C = (L_{max} - L_{min}) / (L_{max} + L_{min})$, where L_{max} and L_{min} represented the maximum and minimum luminance of the stimulus) to 4% (though 10% for the first practice session for an easier task). The first practice session included only one block, which was used for the participants to get familiar with the task. A block contained two interleaved staircases, one for each eye. Each staircase consisted of 60 trials and started with the highest contrast level. The test contrast decreased after two successive correct responses and increased after every wrong response. The step size for the staircase was initially three contrast levels and was reduced to one contrast level after three reversals. The procedure was similar for the other two practice sessions except that each session contained two blocks. Contrast threshold from each staircase was calculated by averaging the contrast levels of the last six reversals. The mean threshold of the latter two practice sessions was used to determine the seven contrast levels in the formal experiments, which were designed with the constant stimuli method. Similar to the formal experiments, feedback beeps (1200 Hz) were given after the response. However, they were delivered randomly in half of the trials with correct responses and participants were told to ignore the beep.

After the practice, participants finished a pre-test session, two training sessions, and a post-test session. The task was same as that used in the practice experiment. Seven test contrast levels ranged logarithmically from one fourth to three times of the threshold estimated in the practice sessions were used. For each eye, every contrast level was tested for 50 times, resulting in 700 trials per session. The test eye and contrast were randomly selected in each trial. A session was divided into 4 blocks, allowing the participants to take a break after every block. The contrast threshold of each eye was estimated by fitting the accuracies at all contrast levels with a Weibull function (82% correct performance).

Unbeknown to the participants, one eye was selected to be the rewarded eye before the formal experiments. The selection of rewarded eye was counter-balanced across participants. Every correct response for a rewarding trial was accompanied with an auditory feedback (1200 Hz). However, only in the training sessions, participants were informed that the high frequency beep meant an extra monetary reward of 0.77 yuan. After each training block, a message on the screen showed the participants how much they had earned.

Experiment 2c. The stimuli were gratings (3° in diameter, 1.5 cpd) with the contrast of 80%, and the orientations were about 45°. In a trial, two test gratings were presented to one of the eyes successively; each grating was presented for 100 ms with a 300-ms gap between them (Figure 2D). Participants performed an orientation discrimination task by judging whether the second grating tilted clockwise or counter-clockwise to the first one.

After a few practice sessions, participants performed a pilot session where their orientation discrimination thresholds were measured. The orientation difference between the reference (45°) and test gratings was adjusted according to a staircase procedure. Each practice session included two interleaved staircases, one for each eye. The pilot session included four interleaved staircases, two for each eye. Every staircase contained 50 trials. Fifty levels of the orientation offset were predetermined for the staircase, ranging logarithmically from 0.1° to 10°. The mean orientation offset from the last six reversals of each staircase was calculated as the orientation discrimination threshold for each eye, respectively (71% correction threshold).

In the formal experiments, participants finished a pre-test session and 3 training sessions. Every session consisted of 5 blocks (100 trials per block). For each eye, four subthreshold offset levels (0.10, 0.23, 0.37, 0.5 × individual threshold) and a threshold offset level were used. The threshold level was set to ensure participants could discriminate the orientation difference in some of the trials so that they would not give up on the task. In each block, 10 trials were tested for each offset level.

There was no reward in the pre-test. In the training sessions, one eye was selected as the rewarded eye. A beep (1300 Hz) would sound immediately if the participants made a correct response to a stimulus with threshold orientation offset but only when it was presented to the rewarded eye. For stimuli with subthreshold orientation offset that were presented to the rewarded eye, the beep was given regardless of whether the responses were correct or not. Participants were not informed under what circumstance the reward would be given, but were instructed that each beep meant that they had earned a certain amount of reward (0.09 yuan). The total gain was presented on the screen after each block. It should be noted that, since participants were rewarded according to the correction of their responses for only 10 trials in each block, the amount of reward could be nearly equal across the blocks. To increase the variance of reward over blocks, a random value was either added to or subtracted from the actual money the participants had earned.

Results and discussion

Experiment 2a

We first tested the role of inter-ocular suppression using trials with and without CFS. Only in the trials without CFS (i.e., target-only trials) were participants rewarded for a correct response to the target presented to the rewarded eye. It was clear that these target-only trials were irrelevant to inter-ocular suppression. We found that in these trials, participants performed well in the pre-test for both eyes, with no significant difference in the performance ($t(14) = .91, p = .377, d = .19$). Training improved the performance for both eyes slightly by between 1.0% and 1.5% in the target-only trials with no difference across the eyes (all FDR corrected ps $> .69$). In the with-CFS trials, the breakthrough ratio increased with training. However, no significant difference was observed between the breakthrough ratio for the rewarded eye and that for the non-rewarded eye in the pre-test ($t(14) = .97, p = .349, d = .22$), and the increase of breakthrough ratios did not show any difference between the two eyes in the training sessions (see Figure 2A, training 1: $t(14) = 1.01$, $p_{\text{uncorrected}} = .331$, $d = .29$; training 2: $t(14) = .47$, $p_{\text{uncorrected}} = .645$, $d = .10$; training 3: $t(14) = .47$, $p_{\text{uncorrected}} = .647$, $d = .10$; and all $p_{\text{FDRcorrected}} > .64$). Similar results were found after the false alarm correction.

Since the performance for the target-only trials was almost perfect, a potential ceiling effect could not be excluded. We then used a more difficult contrast detection task and an orientation discrimination task to further examine this issue. Subthreshold stimuli rather than inter-ocular suppression were used to render the stimuli hard to perceive in these two experiments.

Experiment 2b

In the contrast detection experiment, no significant difference of the performance between the two eyes was observed (Figure 2C, ($F(1, 13) = 1.99, p = .182, \eta^2 = .13$), though the contrast threshold showed a significant change across sessions ($F(1.19, 15.45) = 7.98, p = .010, \eta^2 = .38$, Greenhouse–Geisser corrected). The main effect of session was predominantly due to the increase of thresholds in the post-test than in the other sessions, probably reflecting reduced motivation in the post-test that lacked incentives as compared to the training sessions.

Experiment 2c

Considering that the contrasts of stimuli in Experiment 2b were close to or below the detection threshold, visual signals to primary visual cortex might be faint. As a result, the eye-specific reward might not be able to enhance these signals. However, the eye-specific reward learning effect was still absent in Experiment 2c where high contrast gratings and orientation discrimination task were used. For each offset level, the performances in the pre-test were not statistically different between the two eyes (all $p_{\text{FDRcorrected}} > .80$). After subtracting the correction rates of the pre-test from those in training sessions, no difference was found between the rewarded and non-rewarded eyes on any levels in all the training sessions (Figure 2E, all $p_{\text{FDRcorrected}} > .19$).

Experiment 3: Reward learning of both consciously accessible and consciously inaccessible features

The results of experiments 1 and 2 indicated that inter-ocular suppression was necessary for eliciting the eye-specific reward learning effects when participants could not discriminate the rewarding vs. non-rewarding targets. In Experiment 1a, the eye-of-origin information was the only difference between the two kinds of targets, though was consciously inaccessible. An interesting question is whether we can still observe the eye-specific learning effects when reward is also related to another feature that could be consciously accessible. This issue was investigated in the following experiments.

Method

Participants

Eighty-three volunteers were recruited for Experiment 3 and 34 of them passed the screen test. Eighteen (12 males and 6 females) participated in Experiment 3a, and 16 (5 males and 11 females) participated in Experiment 3b. The participants ranged in age from 18 to 27 years.

Stimuli and procedures

The stimuli were similar to those in Experiment 1a, except that the target was a vertical or horizontal bar (length: 1.4°, linewidth: 0.2°) without the square frame (see Figure 3A). This could make a vertical bar more distinguishable from a horizontal bar.

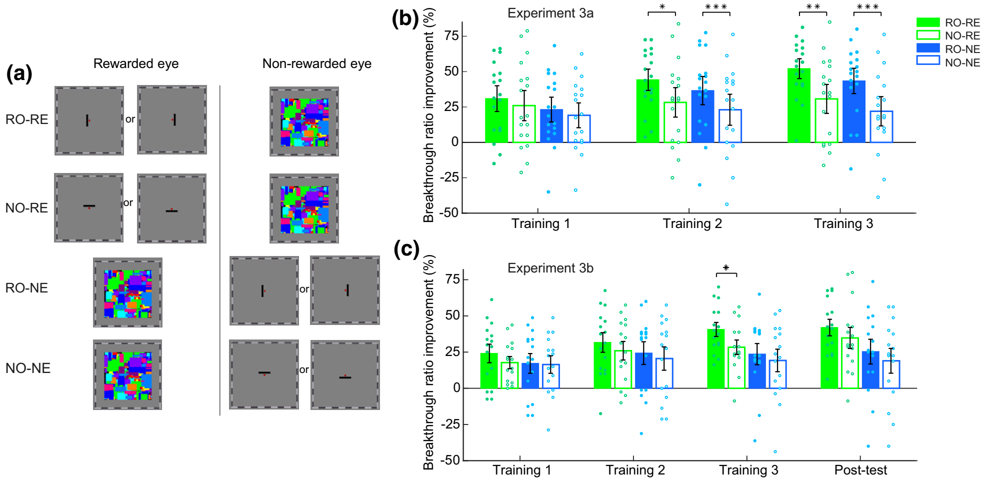

Figure 3. Stimuli and results of Experiment 3. (A) The examples of four kinds of trials: rewarded orientation in the rewarded eye (RO-RE), non-rewarded orientation in the rewarded eye (NO-RE), rewarded orientation in the non-rewarded eye (RO-NE), and non-rewarded orientation in the non-rewarded eye (NO-NE). The selection of rewarded orientation and rewarded eye were counter-balanced across participants. (B)-(C) The improvement of breakthrough ratio relative to pre-test in training and post-test sessions of Experiment 3a and Experiment 3b. Asterisks indicate the significance level (with $^{*}p_{\text{FDRcorrected}} < .05$, $^{**}p_{\text{FDRcorrected}} < .01$, and $^{***}p_{\text{FDRcorrected}} < .001$).

Like Experiment 1a, participants were required to detect a target in one eye suppressed by the CFS stimuli in the other eye. Because there were two different targets (vertical or horizontal) and the target could be presented to one of the two eyes, there were four conditions in this experiment (vertical target in the left eye, vertical target in the right eye, horizontal target in the left eye, and horizontal target in the right eye). Only one of the four conditions (counter-balanced across the participants) was assigned to be the rewarding condition, while the other three conditions were non-rewarding conditions. A correct response in a rewarding trial would produce a reward of 0.5 yuan accompanied by an auditory feedback. The gross of rewards was listed in a message on the screen after the end of each block. The four types of trials were randomly interleaved within a session.

Since the eye-specific learning effect in Experiment 1a was no longer observed once the reward was withdrew in the post-test, we first asked a group of participants to complete the experiment with a screen test, a pre-test, and three training sessions (Experiment 3a). Another group of participants then completed the experiment with an extra post-test without reward after training (Experiment 3b) to test the persistence of the learning effect.

Results and discussion

In this Experiment, the rewarding target was defined by a conjunction of two features. Only the bar in one of the two orientations presented to the rewarded eye was the rewarding target (Figure 3A). We conducted Experiment 3 on two different groups of participants. Because of a mistake, the first group of participants did not complete the post-test after training (this is referred to as Experiment 3a hereafter). Therefore, we replicated the experiment in another group of participants but added the post-test (this is referred to as Experiment 3b hereafter). No significant differences among conditions were found in the pre-test of both groups (all *p*s > .18). Results from the first group of participants showed a significant interaction between orientation and session (see Table S1 for detailed statistics). Paired *t*-test on the improvement of breakthrough ratio revealed significant learning effects in the 2nd and 3rd training sessions, and the effects were specific to the rewarded orientation but not specific to the rewarded eye (Figure 3B, rewarded orientation – rewarded eye (RO-RE) vs. non-rewarded orientation – rewarded eye (NO-RE): training 1: $t(17) = 1.34$, $p_{\text{uncorrected}} = .199$, $p_{\text{FDRcorrected}} = .199$, $d = .18$; training 2: $t(17) = 2.69$, $p_{\text{uncorrected}} = .016$, $p_{\text{FDRcorrected}} = .023$, $d = .62$; and training 3: $t(17) = 4.20$, $p_{\text{uncorrected}} < .001$, $p_{\text{FDRcorrected}} = .002$, $d = .86$; and rewarded orientation – non-rewarded eye (RO-NE) vs. non-rewarded orientation – non-rewarded eye (NO-NE): training 1: $t(17) = 1.56$, $p_{\text{uncorrected}} = .138$, $p_{\text{FDRcorrected}} = .138$, $d = .16$; training 2: $t(17) = 5.09$, $p_{\text{uncorrected}} < .001$, $p_{\text{FDRcorrected}} < .001$, $d = .46$; and training 3: $t(17) = 6.86$, $p_{\text{uncorrected}} < .001$, $p_{\text{FDRcorrected}} < .001$, $d = .78$. All $p_{\text{FDRcorrected}} > .22$ for RO-RE vs. RO-NE and NO-RE vs. NO-NE). Results from the second group of participants showed a more sluggish learning effect (Figure 3C). Repeated measurements ANOVA disclosed a non-significant trend of interaction between orientation and session. An orientation-specific effect was only observed in the last training session in the rewarded eye and was absent in the post-test (RO-RE vs. NO-RE: training 1: $t(15) = 1.50$, $p_{\text{uncorrected}} = .153$, $p_{\text{FDRcorrected}} = .205$, $d = .35$; training 2: $t(15) = 1.31$, $p_{\text{uncorrected}} = .209$, $p_{\text{FDRcorrected}} = .209$, $d = .26$; training 3: $t(15) = 3.33$, $p_{\text{uncorrected}} = .005$, $p_{\text{FDRcorrected}} = .018$, $d = .75$; and post-test: $t(15) = 1.83$, $p_{\text{uncorrected}} = .087$, $p_{\text{FDRcorrected}} = .173$, $d = .33$; and RO-NE vs. NO-NE: all $p_{\text{FDRcorrected}} > .35$). Though there was also a significant interaction between eye of origin and session, none of the difference between two eyes survived the FDR correction (RO-RE vs. RO-NE: training 1: $t(15) = 1.31$, $p_{\text{uncorrected}} = .211$, $p_{\text{FDRcorrected}} = .281$, $d = .31$; training 2: $t(15) = .96$, $p_{\text{uncorrected}} = .352$, $p_{\text{FDRcorrected}} = .352$, $d = .30$; training 3: $t(15) = 2.36$, $p_{\text{uncorrected}} = .032$, $p_{\text{FDRcorrected}} = .065$, $d = .83$; and post-test: $t(15) = 2.59$, $p_{\text{uncorrected}} = .021$, $p_{\text{FDRcorrected}} = .065$, $d = .69$. NO-RE vs. NO-NE: all $p_{\text{FDRcorrected}} > .24$).

Despite that the learning patterns of the two groups were not exactly the same, the learning effects of both groups were more orientation-specific and developed more slowly as compared to that in Experiment 1a. No significant learning effect was observed in the first training session. Similar results were found after false alarm correction. The results suggested that when the reward was associated with the conjunction of a consciously accessible feature and a consciously inaccessible feature, the reward learning depended mainly on the consciously accessible one.

Experiment 4: Testing whether the rewarding target can be consciously differentiated

Though it has been reported that observers were unable to discriminate in which eye the stimuli were presented when viewing monocular patterns (Wolfe & Franzel, 1988; Zhang et al., 2012), some studies have found that the accuracy might be above the chance level if observers were asked to discriminate the eye of origin of the stimuli (Schwarzkopf, Schindler, & Rees, 2010). We thus ran an experiment to investigate whether the participants could discern the rewarding targets in our paradigm.

Method

Participants

Ninety-six participants (38 males and 58 females, 5 have participated in Experiment 1 or 3) finished the test. The participants ranged in age from 18 to 29 years.

Stimuli and procedures

The stimuli were the same as those in Experiment 1a. To test as many participants as possible, we did not strictly screen participants by perceptual eye dominance. Thus, the majority of the recruited participants were allowed to finish the experiments except for three with extreme inter-ocular imbalance (i.e., the breakthrough ratio was always 100% in one eye and 0% in the other eye). We had participants perform a pre-test and a training session first. Each session included one block of 160 trials. Afterwards, participants finished a questionnaire that contained 14 questions (see the Questionnaire at the end of the Supplementary materials). In the questionnaire, participants were asked whether the reward delivery followed some rules and whether the reward was associated with any feature(s) of targets, such as the position, eye of origin, and orientation of targets.

Results and discussion

Since we did not screen participants for relatively balanced ocular dominance that could potentially avoid any ceiling or floor effect of the training-induced change of breakthrough ratio, we did not focus the analysis on the results of b-CFS task.

The results of questionnaire showed that 34 of the 96 participants answered that the reward was associated with the eye of origin of the target. However, 15 of them reported not finding any rules of reward delivery until they were hinted the association between reward and target features (Table S7). Nevertheless, most of them also thought that the reward delivery depended on the target features including the location, the timing, or the orientation of target presentation. Even though there was no relation between reward and those features other than the eye of origin of the target and the questionnaire provided the option 'The feature was not associated with reward', these participants pretended to know that the reward was associated with some of the features. Further analysis on the responses of these participants showed that 21 of them correctly recognized the rewarded eye, but only 2 discerned that the reward delivery was exclusively associated with the eye of origin of target (Figure 4). Thus, we believe that most of the participants in our study were unable to identify the reward-associated targets during the training experiments.

General discussion

We found an eye-specific learning effect by training participants with a b-CFS reward learning paradigm where monetary rewarding and non-rewarding cues were rendered identical to each other except for their eye-of-origin information. Our design rests on the phenomenon that when a monocular pattern is presented to one eye, people have no explicit knowledge of the pattern's eye of origin (Wolfe & Franzel, 1988; Zhang et al., 2012). Thus, we could examine whether the reward coding system tightly depends on the conscious discerning of reward-associated stimuli.

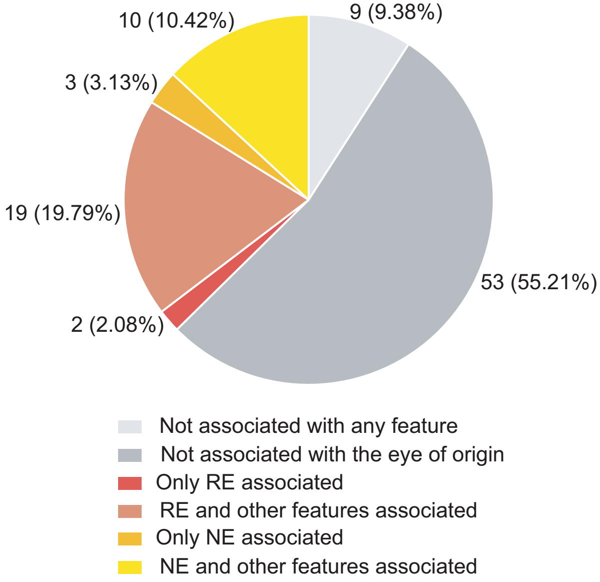

Figure 4. The distribution of numbers of participants' responses to the association between reward and the features of target. The proportions were displayed in the brackets.

The finding of an eye-specific learning effect in Experiment 1a supports our hypothesis that the reward coding system can rely merely on the eye-of-origin information of the reward-associated stimuli to produce an unconscious reward-based learning effect. This result differs from the findings in a well-known perceptual learning study (Seitz, Kim, & Watanabe, 2009). First, the reward was water for thirsty observers in their experiment rather than money here. Second, their rewarding and non-rewarding stimuli had different orientations, whereas the only feature difference between our stimuli (Experiment 1a) was eye of origin. Third and most important, their learning effect developed over 20 days of training and could be observed in a separate sensitivity test where no reward was given, whereas ours established quickly during the training and vanished immediately in a post-test without reward conducted shortly after the training. These distinct characteristics suggest that Seitz et al.'s findings, as they also proposed (Seitz et al., 2009), should be considered as a type of perceptual learning effect. By contrast, the present findings remind us of the brevity of non-conscious fear conditioning (Raio, Carmel, Carrasco, & Phelps, 2012). However, our rapid non-conscious reward learning effect is closer to an operant learning effect, which was induced by monetary reward, a positive and pleasant rather than a negative and threatening event (e.g., fear). Importantly, the stimulus-reward pairing was established only based on eye of origin – a consciously inaccessible feature.

In Raio et al., (2012)'s non-conscious fear conditioning study, the non-conscious fear learning was only significant during early acquisition and declined quickly in the second half of trials. Nevertheless, the non-conscious reward learning in our Experiment 1a did not show a rapid forgetting during training. This might be due to different attentional status during the training in the two studies. Raio et al., (2012) used a CFS paradigm in which participants did not have a task related to the suppressed stimuli, whereas we used a b-CFS paradigm so that participants had to pay attention and report any potential breakthroughs of the suppressed stimuli. Not only that, we even found an important role of attention in generating the eye-specific reward-based learning effect, because the effect was absent when attention was distracted from the rewarding task. Notably, the contribution of top-down attention here should differ clearly from the more common roles of selective attention in modulating the actions or perceptions when participants have an explicit knowledge of the reward-associated feature like in Experiment 3 and many previous studies (Marx & Einhauser, 2015; Raymond & O'Brien, 2009; Wilbertz et al., 2014).

The current work provides a strong case arguing that reward can induce learning in human V1, which is largely independent of consciousness. Such relatively lower level learning effect should be driven by the co-work of top-down eye-based attention (Zhang et al., 2012) and diffusely distributed neuromodulatory signals released by the reward coding system (Dalley et al., 2001; Schultz, 2000; Vickery et al., 2011). During inter-ocular suppression, the invisible rewarding target activates monocular neurons for the rewarded eye, while the activities of binocular neurons are mainly dominated by the signals for the CFS stimuli. Obviously, only the firings of monocular neurons for the rewarded eye are highly predictive of later rewards. During the training, the reward coding system may soon detect the reliable association between rewards and responses of monocular neurons for the rewarded eye. Eye-based attention may then increase the gains particularly for those neurons. As a result, the breakthrough was facilitated more for the rewarded eye than for the non-rewarded eye in Experiment 1a. Once the attentional resources were consumed by another demanding task, the co-work of attention and reward coding system failed; thus, no eye-specific learning effect was observed in Experiment 1b.

The above explanation receives further support from the results of our Experiment 2 which examined the critical role of inter-ocular suppression. Without inter-ocular suppression, the targets were represented by the activities of both monocular and binocular neurons. In case of breakthrough, the firings of monocular neurons were either 100% (for the rewarded eye) or 0% (for the non-rewarded eye) predictive of subsequent rewards, yet the firings of binocular neurons were always 50% predictive of rewards. The absence of the eye-specific learning effects thus indicated that in Experiment 2, the reward coding system weighted heavily on the activities of binocular neurons and ignored the eye-of-origin information. This is possible given that binocular neurons greatly outnumber monocular neurons in the visual cortex.

As indicated by Experiment 3, inter-ocular suppression is not sufficient to the eye-specific learning effects. Although both orientation and eye-of-origin information provided rewarding cues, only the orientation difference of the rewarding and non-rewarding targets, but not the eye-of-origin difference, was consciously discerned. Given the easily identified orientation difference, the reward coding system might treat both RO-RE and RO-NE targets as a single type of targets that sometimes (50% probability) brought rewards at the time of breaking into awareness. As a result, the reward coding system may selectively strengthen the representations in the expected orientation, just as its role in teaching attention to make selections (Marx & Einhauser, 2015; Raymond & O'Brien, 2009; Wilbertz et al., 2014).

As compared to Experiment 1a, participants in Experiment 3 learned more slowly. The more complicated reward rule in Experiment 3 might hamper a fast reward-based learning. Alternatively, the sluggish learning might result from the additional use of a consciously accessible feature (i.e., orientation), whereas in Experiment 1a, only the consciously inaccessible feature was used to associate with rewards. This also agrees with Raio et al., (2012)'s finding that participants learned more slowly when they were aware of the reward-associated stimuli than when not.

In Experiment 4, we examined whether participants were able to consciously differentiate the rewarding and non-rewarding targets. Earlier work chose objective methods (e.g., two-alternative forced-choice (2AFC)) to measure consciousness, because subjective methods are considered not bias-free. However, recent consciousness work has warned that objective methods may overestimate conscious knowledge – above-chance performance on a 2AFC task may be due to unconscious knowledge rather than conscious (Timmermans & Cleeremans, 2015). Moreover, objective measurements request a closer inspection of the stimuli. Since the more salient CFS stimuli dominated the awareness for most of the time in our paradigm, attention would be directed considerably more to the CFS stimuli in a 2AFC task than during the training (where the CFS stimuli were task- or motivation-irrelevant). Such change of attentional state has been found to result in overestimated discriminability (Vermeiren & Cleeremans, 2012). Considering these obvious limitations of objective measurements, we chose the questionnaire investigation. If participants had any conscious knowledge of the difference between the rewarding and non-rewarding targets, it would certainly be beneficial to further this sense of consciousness onto a self-monitoring stage (Dehaene, Lau, & Kouider, 2017) so that this knowledge could be kept in working memory and the reward coding system could constantly use this knowledge to pursue more monetary rewards. Then, it would be very likely that they could correctly report the difference between the two kinds of targets. However, only 2 out of 96 participants correctly reported that the reward delivery was exclusively associated with the eye of origin of the target. Most participants even reported that reward was associated with other target features. Additional evidence is from

Experiment 3 where the rewarding cue was defined as a conjunction of orientation and eye of origin. The slower development of an orientation-specific learning effect confirmed that participants were indeed motivated to remember and use any conscious knowledge that they could grasp. However, no reliable eye-specific learning effect was observed in that experiment. Since conscious knowledge is knowledge one can control the use of (Fu, Fu, & Dienes, 2008; Jacoby, 1991; Timmermans & Cleeremans, 2015), we believe that most participants had no conscious knowledge about the difference between the rewarding and non-rewarding targets in this experiment and Experiment 1.

We therefore propose that the reward coding system can produce two different types of reward-based learning. One of them is independent of the conscious discrimination of the rewarding stimuli yet fairly consuming attentional resources, likely occurring as early as in V1. We call it the unconscious reward learning effect. The other type of learning rests on the close interactions between reward and selective attention, which we call the conscious reward learning effect. The conscious reward learning effect can overshadow the unconscious reward learning effect when voluntary attention begins to select stimuli of behavioural significance. The eye-specific unconscious reward learning effect is likely to develop very fast in the training and highly depends on the context of reward delivery. As soon as the context of reward delivery was absent (e.g., in the post-test), the effect was also absent. Once the conscious reward learning joined, the temporal pattern of learning changed, hinting distinct timescales between the two types of reward-based learning.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (31871104, 31571112, 31500888, and 31830037), the Strategic Priority Research Program and the Key Research Program of Chinese Academy of Sciences (XDB32010300 and QYZDB-SSW-SMC030), Beijing Municipal Science & Technology Commission, and the Fundamental Research Funds for the Central Universities.

Conflict of interest

All authors declare no conflict of interest.

Author contributions

Yi Jiang (Funding acquisition; Methodology) Min Bao (Conceptualization; Funding acquisition; Methodology; Resources; Supervision; Writing – original draft; Writing – review & editing) Bo Dong (Investigation) Xue Dong (Formal analysis; Investigation; Methodology; Software; Writing – original draft; Writing – review & editing) Mingxia Zhang (Formal analysis; Funding acquisition; Investigation; Software).

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011a). Learned value magnifies salience-based attentional capture. *PLoS One*, *6*, e27926. https://doi.org/10.1371/journal.pone.0027926

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011b). Value-driven attentional capture. *Proceedings of the National Academy of Sciences USA*, *108*(25), 10367–10371. https://doi.org/10.1073/pnas.1104047108

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436. https://doi.org/10.1163/156856897X00357

Dalley, J. W., McGaughy, J., O'Connell, M. T., Cardinal, R. N., Levita, L., & Robbins, T. W. (2001). Distinct changes in cortical acetylcholine and noradrenaline efflux during contingent and noncontingent performance of a visual attentional task. *Journal of Neuroscience*, *21*, 4908–4914. https://doi.org/10.1523/JNEUROSCI.21-13-04908.2001

Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness, and could machines have it? *Science*, *358*, 486–492. https://doi.org/10.1126/science.aan8871

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fu, Q., Fu, X., & Dienes, Z. (2008). Implicit sequence learning and conscious awareness. *Conousness & Cognition*, *17*(1), 185–202. https://doi.org/10.1016/j.concog.2007.01.007

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience*, *30*(33), 11096–11103. https://doi.org/10.1523/JNEUROSCI.1026-10.2010

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, *30*, 513–541. https://doi.org/10.1016/0749-596X(91)90025-F

Jiang, Y., Costello, P., & He, S. (2010). Processing of invisible stimuli: Advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological Science*, *18*, 349–355. https://doi.org/10.1111/j.1467-9280.2007.01902.x

Marx, S., & Einhauser, W. (2015). Reward modulates perception in binocular rivalry. *Journal of Vision*, *15*(1), 11. https://doi.org/10.1167/15.1.11

Mastropasqua, T., Tse, P. U., & Turatto, M. (2015). Learning of monocular information facilitates breakthrough to awareness during interocular suppression. *Atten Percept Psychophys*, *77*, 790–803. https://doi.org/10.3758/s13414-015-0839-z

Proshansky, H., & Murphy, G. (1942). The effects of reward and punishment on perception. *Journal of Psychology: Interdisciplinary and Applied*, *13*, 295–305. https://doi.org/10.1080/00223980.1942.9917097

Raio, C. M., Carmel, D., Carrasco, M., & Phelps, E. A. (2012). Nonconscious fear is quickly acquired but swiftly forgotten. *Current Biology*, *22*, R477–479. https://doi.org/10.1016/j.cub.2012.04.023

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. https://doi.org/10.1111/j.1467-9280.2009.02391.x

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience*, *1*, 199–207. https://doi.org/10.1038/35044563

Schwarzkopf, D. S., Schindler, A., & Rees, G. (2010). Knowing with which eye we see: Utrocular discrimination and eye-specific signals in human visual cortex. *PLoS One*, *5*(10), 1–8. https://doi.org/10.1371/journal.pone.0013775

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, *61*, 700–707. https://doi.org/10.1016/j.neuron.2009.01.016

Thorndike, E. L. (1911). *Animal intelligence*. New York: Macmillan.

Timmermans, B., & Cleeremans, A. (2015). *How can we measure awareness? An overview of current methods*. In (Vol. 21). Oxford: Oxford University Press.

Vermeiren, A., & Cleeremans, A. (2012). The validity of d′ measures. *PLoS One*, *7*(2), e31595. https://doi.org/10.1371/journal.pone.0031595

Vickery, T. J., Chun, M. M., & Lee, D. (2011). Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron*, *72*(1), 166–177. https://doi.org/10.1016/j.neuron.2011.08.011

Wilbertz, G., van Slooten, J., & Sterzer, P. (2014). Reinforcement of perceptual inference: reward and punishment alter conscious visual perception during binocular rivalry. *Frontiers in Psychology*, *5*, 1377. https://doi.org/10.3389/fpsyg.2014.01377

Wolfe, J. M., & Franzel, S. L. (1988). Binocularity and visual search. *Percept Psychophys*, *44*(1), 81–93. https://doi.org/10.3758/BF03207480

Xue, X., Zhou, X., & Li, S. (2015). Unconscious reward facilitates motion perceptual learning. *Visual Cognition*, *23*, 1–18. https://doi.org/10.1080/13506285.2014.981625

Zhang, P., Jiang, Y., & He, S. (2012). Voluntary attention modulates processing of eye-specific visual information. *Psychological Science*, *23*(3), 254–260. https://doi.org/10.1177/0956797611424289

Received 9 December 2020; revised version received 8 May 2021

Supporting Information

The following supporting information may be found in the online edition of the article:

Table S1. Statistics for the results of repeated measurements ANOVA.
Table S2. Finally determined target contrasts and the breakthrough ratios at the determined contrasts for participants who completed the whole procedure in Experiment 1a.
Table S3. Finally determined target contrasts and the breakthrough ratios at the determined contrasts for participants who completed the whole procedure in Experiment 1b.
Table S4. Finally determined target contrasts and the breakthrough ratios at the determined contrasts for participants who completed the whole procedure in Experiment 2a.
Table S5. Finally determined target contrasts and the breakthrough ratios at the determined contrasts for participants who completed the whole procedure in Experiment 3a.
Table S6. Finally determined target contrasts and the breakthrough ratios at the determined contrasts for participants who completed the whole procedure in Experiment 3b.
Table S7. Answers to five questions from participants who thought there was an association between reward and eye of origin of target (34/96).