

Unconscious reward facilitates motion perceptual learning

Xin Xue^{1,2,3}, Xiaolin Zhou^{1,2,3,4}, and Sheng Li^{1,2,3}

¹Department of Psychology, Peking University, Beijing, China

²Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China

³PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China

⁴Center for Brain and Cognitive Sciences, Peking University, Beijing, China

(Received 1 June 2014; accepted 23 October 2014)

Reward signal plays an important role in guiding human learning behaviour. Recent studies have provided evidence that reward signal modulates perceptual learning of basic visual features. Typically, the reward effects on perceptual learning were accompanied with consciously presented reward during the learning process. However, whether an unconsciously presented reward signal that minimizes the contribution of attentional and motivational factors can facilitate perceptual learning remains less well understood. We trained human subjects on a visual motion detection task and subliminally delivered a monetary reward for correct response during the training. The results showed significantly larger learning effect for high reward-associated motion direction than low reward-associated motion direction. Importantly, subjects could neither discriminate the relative values of the subliminal monetary reward nor correctly report the reward-direction contingencies. Our findings suggest that reward signal plays an important modulatory role in perceptual learning even if the magnitude of the reward was not consciously perceived.

Keywords: Reward; Perceptual learning; Unconscious; Reinforcement learning; Feedback.

Perceptual learning refers to the phenomenon that subjects could obtain long-lasting improvement in perceptual sensitivity resulting from visual perceptual experience after their critical period of plasticity (Gilbert, Sigman, & Crist, 2001; Sagi & Tanne, 1994). Previous research on perceptual learning mainly focused on the cortical loci where the learning effects took place. Whereas early literatures

Please address all correspondence to Sheng Li, Department of Psychology, Peking University, 5 Yiheyuan Road, Haidian, Beijing, 10087, China. E-mail: sli@pku.edu.cn

suggested training-induced plastic changes at early sensory cortex (Karni & Sagi, 1991; Schoups, Vogels, Qian, & Orban, 2001), later investigations suggested that high-order mechanisms involved decision making can account for the effects observed in perceptual learning (Doshier & Lu, 1998, 1999; Law & Gold, 2009; Xiao et al., 2008; Zhang et al., 2010). Meanwhile, another important issue for perceptual learning, the gating factor for plasticity, has been widely addressed (Ahissar & Hochstein, 1993; Roelfsema, van Ooyen, & Watanabe, 2010; Sasaki, Nanez, & Watanabe, 2010). Given the importance of attention in sensory and perceptual processing, it is not surprised that attention was indicated as a major factor for gating the learning effect (Ahissar & Hochstein, 1993). Interestingly, reinforcement signals have also been suggested to foster perceptual learning in both task-relevant and task-irrelevant manners (Frankó, Seitz, & Vogels, 2009; Kahnt, Grueschow, Speck, & Haynes, 2011; Law & Gold, 2009; Seitz & Watanabe, 2003; Seitz, Kim, & Watanabe, 2009; Seitz, Lefebvre, Watanabe, & Jolicoeur, 2005). Specifically, Seitz and Watanabe (2005) proposed that both task-relevant and task-irrelevant perceptual learning result from interactions between bottom-up sensory signal and concurrent reinforcement signal. As a special case, a task-irrelevant feature can be learned when it is presented temporally coincident with the reinforcement signal. A task-relevant feature, on the other hand, is by itself presented concurrently with the reinforcement signal of task engagement and thus naturally causes learning to occur. This theoretical framework was supported by both empirical and computational studies (Kahnt et al., 2011; Law & Gold, 2009; Seitz et al., 2005, 2009).

However, a critical issue yet to be addressed concerning the role of reward signal in perceptual learning is whether unconsciously presented reward can facilitate learning. This is important because such unconsciously presented reward stimulus will allow us to isolate the reward-induced learning effect by minimizing the contribution from attentional and motivational factors (Maunsell, 2004). Previous study has shown that attention can modulate brain activity in the primary visual cortex under binocular suppression (Watanabe et al., 2011). It has also been shown that subjects can learn the associated value or punishment property through a conditioning process driven by a conscious unconditioned stimulus even if the conditioned stimulus is blocked by continuous flash suppression or a sandwich mask from reaching conscious awareness (Pessiglione et al., 2008; Raio, Carmel, Carrasco, & Phelps, 2012). Furthermore, the learned reward association in turn can modulate attention and working memory (Anderson, Laurent, & Yantis, 2011; Della Libera & Chelazzi, 2006, 2009; Gong & Li, 2014; Hickey, Chelazzi, & Theeuwes, 2010, 2011) and subsequently accelerate the detection of a masked visual stimulus (O'Brien & Raymond, 2012). Thus, the delivery of a conscious reward may introduce attentional bias or transient attentional enhancement for the trained visual stimulus during the process of reward-stimulus conditioning.

In the present study, we set out to directly investigate the effect of unconsciously presented reward on perceptual learning in a coherent motion detection task. We used a masked monetary reward to minimize the reward-induced attentional and motivational biases. In Experiment 1, we adopted the paradigm of subliminal reward priming (Bijleveld, Custers, & Aarts, 2009; Pessiglione et al., 2007) to verify the incentive effect of subliminally presented monetary cue on a coherent motion detection task. In Experiment 2, we examined modulatory effect of unconsciously presented monetary reward on perceptual learning of a coherent motion detection task.

EXPERIMENT 1

Participants

Twenty-four participants (10 males; mean age, 21.8 years) participated in Experiment 1. Two participants were excluded from the analysis due to poor behavioural performance (3 *SDs* above the mean for detection threshold). All participants were students from Peking University, had normal or corrected-to-normal vision, and gave informed consent. The experimental procedures were approved by the local ethics committee. Participants were provided monetary compensation based on performance (mean = ¥27.15 ± 1.64).

Apparatus and stimuli

Visual stimuli were displayed on a CRT monitor (refresh rate, 100 Hz; resolution, 1024 × 768; size, 19.6 inches). The participants viewed the monitor from a distance of 60 cm in a dimly-lit room. A red fixation point (0.15° × 0.15°) on a black background was always present at the centre of the monitor during a trial.

We used dynamic random-dot stimuli (Newsome & Pare, 1988), which consisted of 200 anti-aliased round dots (luminance, 63.8 cd/m²; diameter, 0.075°; density, 16.98 dots·deg⁻²·s⁻¹) displayed within a circular aperture (diameter, 10°). A proportion of the dots moved coherently in the 45° direction at a constant speed of 6°s⁻¹. When a dot moved out of the circular aperture, it wrapped around from the other side of the aperture at a new random location.

Images of Chinese coins of 50 cents (¥0.5) and one cent (¥0.01) were used as reward cue. The two coin images were matched in size (2.3° × 2.3°) and luminance (13 cd/m²). We used an orange-and-white checkerboard as forward and backward masks. The checkerboard was the same size and similar in colour to both coins, and was surrounded by a black circle to cover the salient edge of the coins.

Procedure

Participants were required to perform a coherent motion detection task (Figure 1). Random-dot stimuli were displayed in two successive 150 ms intervals spaced by

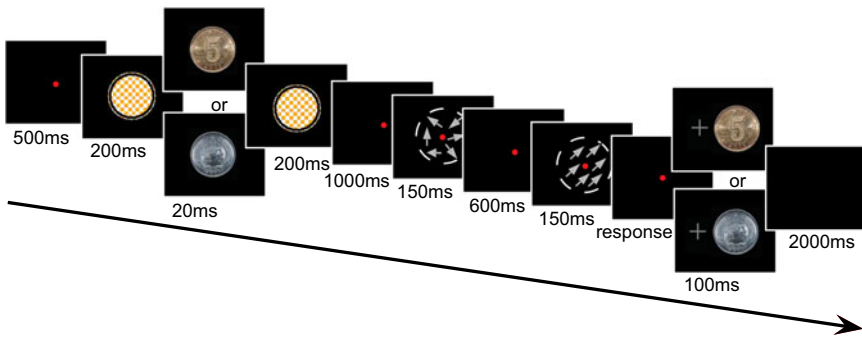


Figure 1. Schematic illustration of the procedure for Experiment 1. For each trial, participants received the monetary reward indicated by the coin image masked by orange-and-white checkerboards if a correct response was made for the motion detection task.

a 600 ms blank period. One of two intervals was embedded with a variable proportion of coherent motion, and the other one was a random motion. Participants were asked to detect which interval contained the coherent motion. Before the presentation of the two random-dot stimuli, an image of 50-cent or one-cent coin was presented for 20 ms as an incentive cue. A sandwich mask, with high contrast and similar in colour to the two coins, was presented for 200 ms immediately before and after the coin image. Participants were informed that they could earn the masked coin by correctly responding in the motion detection task in each trial, but the coin was difficult to perceive. Five levels of signal-to-noise ratio for the coherent motion (5%, 7%, 10%, 14%, and 21%) were measured to obtain psychometric functions of motion detection task for the high reward and low reward conditions respectively. Twelve blocks of 30 trials (360 trials in total) were measured for each participant within a single session. The order of the signal-to-noise ratio and rewarding condition were randomized across blocks. Supraliminal monetary feedback was given at the end of each trial and the cumulative payoff was shown at the end of each block. Participants were informed that their payment was the final cumulative reward divided by 2.5 and adding a basic payment of ¥5. Each participant received between ¥25 and ¥30 as compensation after an hour of experiment.

Before the experiment, 20 trials of 50% signal-to-noise ratio motion were practiced until 90% accuracy was reached. To ensure that participants were familiar with the reward cue images, 60 trials of 50% coherent motion with a preceding supraliminal reward cue of 150 ms were practiced. No real monetary reward was accumulated for the practice trials.

After the experiment, all participants performed a value awareness test of the reward cue using a two-alternative forced choice (2AFC) task to check whether the masked coins were indeed invisible during the experiment. A total number of 80 trials of the sandwich-masked coin (40 trials for 50 cents, and 40 trials for one

cent) were presented in exactly the same way as in the main experiment. Participants had to indicate whether a 50-cent coin or a one-cent coin was presented for each trial. Finally, to verify whether 50 cents and one cent were indeed different reward values to the participants, 13 out of 24 participants were asked to rate the degrees of happiness (seven-point scale, from -3 to 3) for earning 50 cents and one cent, respectively.

Results and discussion

Subliminal incentive motion detection task

As shown in [Figure 2A](#), behavioural data on accuracy were fitted with cumulative Weibull function using maximum-likelihood method to obtain participants' psychometric functions ([Wichmann & Hill, 2001](#)). We compared the

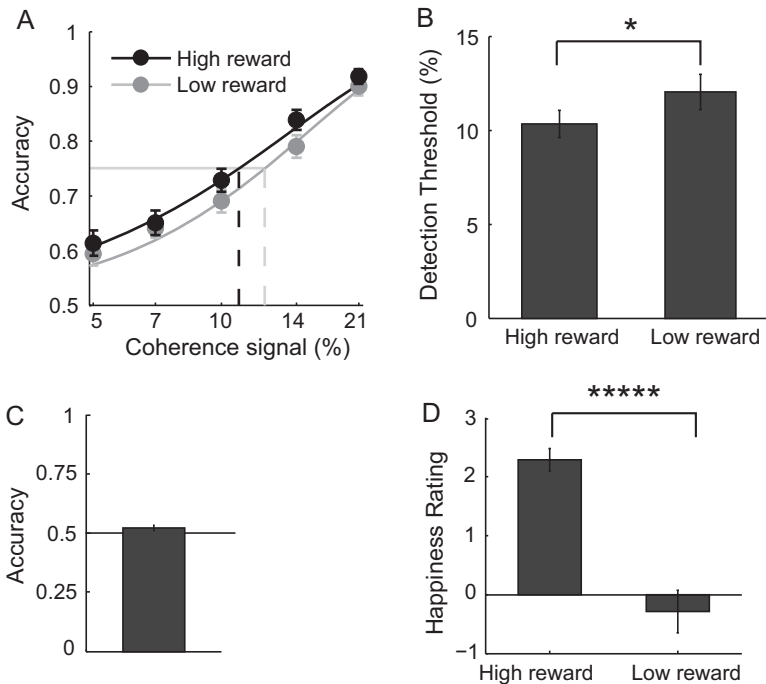


Figure 2. Results of Experiment 1. (A) Psychometric functions fitted for the average accuracy across all participants in the high-reward (black) and low-reward (grey) conditions. (B) Average detection threshold at 75% accuracy across participants under high-reward and low-reward conditions. (C) Chance-level accuracy in the reward awareness test. (D) Happiness ratings for 50 cents versus 1 cent. Error bars represent the standard error of the mean across participants.

detection threshold of the subliminal high reward and low reward conditions to determine whether unconscious reward cue can modulate the performance of coherent motion detection task. Threshold at 75% accuracy was identified from the psychometric function for each participant (Figure 2B). A paired-sample t test showed that the threshold in the high reward condition was significantly lower than that in the low reward condition (mean difference = 1.71%; SE = 0.79%; $t_{21} = -2.15$, $p = .043$).

Reward cue awareness test and reward appraisal test

As indicated by a one-sample t test (Figure 2C), the averaged accuracy across all participants in the reward cue awareness test showed a trend of significance to be above the chance level (mean = 0.522, SE = 0.011; $t_{21} = 1.92$, $p = .07$). But importantly, the two-tailed binomial test on individual accuracies of the reward awareness test showed that no participant reached the significance level (all $p > .15$, except two participants with $p = .09$). Furthermore, none of the participants reported conscious awareness of the reward value in an interview after the experiment. In addition, earning 50 cents was rated happier than one cent (mean difference = 2.57, SE = 0.43; $t_{12} = 5.50$, $p < 10^{-4}$, Figure 2D). These results showed that the participants were largely, if not completely, blocked from reaching conscious awareness of the value of the reward cue, and the value difference was large enough to give them different degrees of pleasure.

Taken together, the results of Experiment 1 demonstrate that unconsciously perceived reward cue can facilitate the detection of coherent motion stimuli. These findings replicated the previous literatures on the unconscious incentive effect for tasks requiring physical effort (Pessiglione et al., 2007), temporal attention (Bijleveld, Custers, & Aarts, 2011), and cognitive engagement (Bijleveld, Custers, & Aarts, 2010; Zedelius, Veling, & Aarts, 2011), and these results validated the usage of masked reward stimuli in Experiment 2.

EXPERIMENT 2

Experiment 2 comprised a pre-test session (day 1), a training session of four days (days 2–5), and a post-test session (day 6) (Figure 3A). The key question was whether unconscious reward feedback during training sessions has an impact on the perceptual learning of coherent motion detection. Therefore, we trained participants on two directions of motion. One direction was associated with high reward and the other one was associated with low reward. If unconsciously presented reward feedback can modulate learning, improvement in the high reward direction would differ from that in the low reward direction.

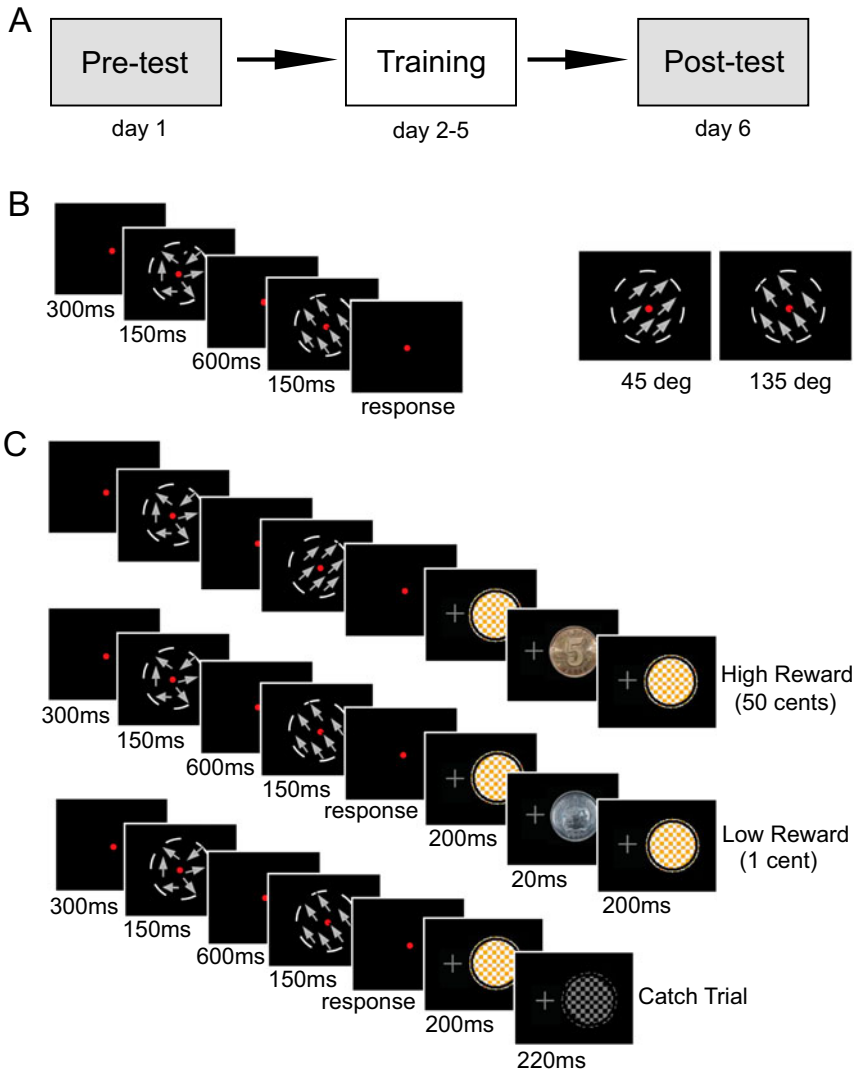


Figure 3. Schematic illustration of the procedure for Experiment 2. (A) General procedure. (B) Procedure of a trial in the pre-test and post-test sessions (left). Dynamic random dots of two directions (45° and 135°) were used in the experiment (right). (C) Procedures of three trial types in the training session: high reward, low reward and catch trials. The masks were the orange-and-white checkerboards, except that a black and grey checkerboard mask was used for the catch-trials.

Participants

Seventeen new participants were recruited from Peking University in Experiment 2 (eight males; mean age: 20.8 ± 1.8 years). All participants had normal or

corrected-to-normal vision, and gave informed consent. The experimental procedures were approved by the local ethics committee. Participants were paid ¥25 as compensation in the pre- and post-test sessions. The payment in the training session was based on their performance (mean: ¥ 28.05 ± 0.34 per day).

Apparatus and stimuli

The apparatus and stimuli were identical to those in Experiment 1, except that a black and grey checkerboard was used as a catch-trial stimulus during the training session.

Procedure

Pre-test and post-test sessions

The task and stimuli in the pre-test and post-test sessions were identical to those in Experiment 1, with the exception that no incentive cue, reward feedback, or error feedback was given (Figure 3B). In the pre- and post-test sessions, 400 trials were measured for each motion direction. The two directions were presented in blocks in an ABAB order, and were counterbalanced between participants. Participants were informed about the motion direction of each block before the block was started.

At the beginning of the pre-test session, participants practiced at least 60 trials of 50% coherent motion until reaching 90% accuracy. In addition, two blocks of 40 trials with the same signal-to-noise ratios as in the test session were practiced, in order to prevent underestimated performance in the pre-test session due to participants' unfamiliarity with the difficulty of the task.

Training session

The task procedure during the training session was the same as in the pre-test and post-test sessions, except that the adaptive procedure of QUEST was used (Watson & Pelli, 1983). The motion coherence of each QUEST started at 40% signal, and updated trial by trial with a Bayesian adaptive method controlling at 75% accuracy. Reward feedback and auditory error feedback were given for each trial (Figure 3C). When the response was correct, a masked image of a 50 cent or one cent coin was presented immediately after the response was made. The reward magnitude was associated with the motion direction. For half of the participants, a correct response on 45° direction always received 50 cents reward subliminally, and only one cent was delivered for the 135° direction. The direction-reward contingency was reversed for the other half of the participants. A black screen and a beep were presented for incorrect response.

Previous studies have shown that spatial and temporal attention is necessary for effective subliminal processing (Dehaene, Changeux, Naccache, Sackur, &

Sergent, 2006; Naccache, Blandin, & Dehaene, 2002; Watanabe et al., 2011). In order to maintain participants' attention on the subliminal reward image, catch trials with an oddball detection task were added to the training session. Participants had to press a third key with the left hand within 1.5 s when a black-and-grey checkerboard, rather than a masked coin, was presented. They were informed that the reward payment was the result of their accuracy in the catch-trial task multiplied by the cumulative money they earned and then divided by 18. We had this manipulation in order to ensure that participants paid attention to the reward feedback across the whole training session. Finally, participants received about ¥8 reward in addition to the ¥20 basic payment on each daily training session.

During the training session, the two motion directions were presented in separate blocks. Participants were informed about the motion direction before the block was started. Twenty-four blocks of 40 trial QUESTs were conducted for each day of the training. QUEST blocks were divided into three segments, in which the order of four blocks of high reward and four blocks of low reward were randomized. Information about the accumulated money was shown when a segment was finished. Therefore, participants had conscious knowledge about their current reward only after finishing equal numbers of blocks of the high reward and low reward directions, eliminating the possibility of inferring the direction-reward contingencies. A total of 1920 trials were completed in each direction after training.

Reward awareness test

Two control tests on reward awareness were conducted to examine (1) whether the reward value was unconscious to the participants during the training session, and (2) whether the reward-direction contingencies were explicitly known by the participants.

In the reward value 2AFC test, accuracy of coin-value discrimination was measured before and after the training session. The stimulus and procedure were exactly the same as in the reward awareness test of Experiment 1. In the reward-direction contingency test after the post-test session, participants were told of the existence of reward and motion-direction contingencies. Participants had to press a key within 3 s to indicate which direction had always associated with higher reward. The limited response time was introduced to avoid inference from related cues, such as the recollection of motion direction with better performance. If the participants could not understand what the question was and ran out of time, they were allowed to answer the question again. After guessing, participants were asked whether they knew there was a motion direction-reward contingency before they were informed. In addition, 15 out of 17 participants

were asked to rate the degrees of happiness (seven-point scale, from -3 to 3) for earning 50 cents and one cent, respectively.

Results and discussion

Motion detection task: pre-test and post-test

Behavioural data on accuracy from the pre-test and post-test sessions were fitted with the Weibull function (Figure 4A). For each participant, the threshold and slope of the psychometric functions at 75% accuracy were identified. Participants' improvement after training was calculated as:

$$\text{Threshold decrease index} = \frac{(\text{pre-test threshold} - \text{post-test threshold})}{\text{pre-test threshold}} \times 100\%. \quad (1)$$

$$\text{Slope increase index} = \frac{(\text{post-test slope} - \text{pre-test slope})}{\text{pre-test slope}} \times 100\%. \quad (2)$$

Paired t -tests showed that the threshold of the high reward direction decreased significantly after training (mean difference = -2.22% , SE = 0.51% ; $t_{16} = 4.35$, $p < .001$) and the low reward direction showed a trend of significance on threshold decrease (mean difference = -1.27% , SE = 0.61% ; $t_{16} = 2.06$, $p = .056$, Figure 4B). Nevertheless, the pre-test thresholds for the two directions did not differ significantly (mean difference = 0.29% , SE = 0.5% ; $t_{16} = 0.59$, $p = .56$), suggesting that the reward effect on the threshold decrease could not be explained by the threshold difference in the pre-test session. Furthermore, a paired t -test on the threshold decrease index showed that the threshold decrease for the high reward direction was significantly larger than that for the low reward direction (mean difference = 13.86% , SE = 6.43% ; $t_{16} = 2.16$, $p = .047$, Figure 4C).

Analysis on the fitted slope showed consistent results with those on the threshold. Paired t -tests showed that the slope for the high reward direction increased after training (mean difference = 2.27% , SE = 1.02% ; $t_{16} = 2.23$, $p = .04$), but the slope for the low reward direction in the post-test session did not differ from that in the pre-test session (mean difference = 0.76% , SE = 0.62% ; $t_{16} = 1.23$, $p = .24$, Figure 4D). Again, the slope learning index for the high reward direction was significantly larger than that in the low reward direction (mean difference = 49.53% , SE = 22.92% ; $t_{16} = 2.16$, $p = .046$, Figure 4E). Again, the slope from the pre-test session for the two directions did not differ significantly (mean difference = -0.33% , SE = 0.98% ; $t_{16} = -0.34$, $p = .74$), suggesting that the reward effect on slope improvement was not due to the slope difference in the pre-test session.

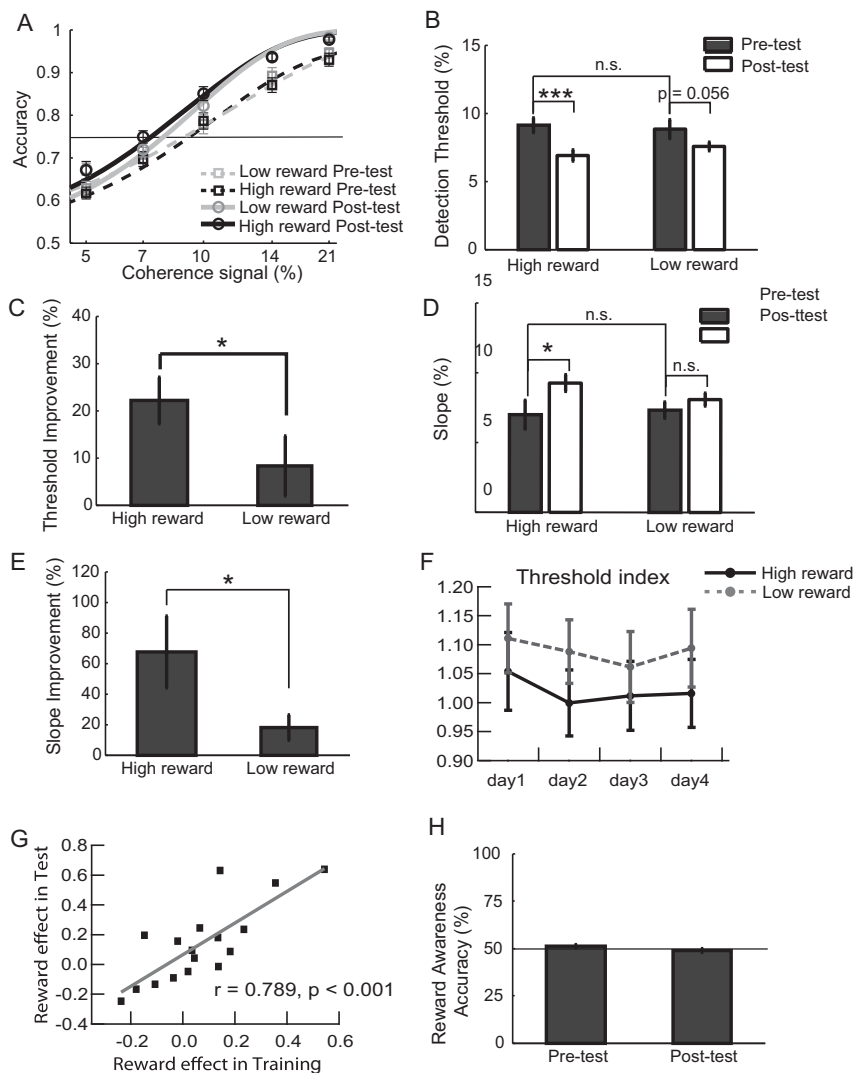


Figure 4. Results of Experiment 2. (A) Psychometric functions fitted for the average accuracy across participants in the high reward (black) and low reward (grey) conditions for the pre-test (dashed line) and the post-test (solid line) sessions. (B) Average detection threshold at 75% accuracy across participants under the high reward and low reward conditions for the pre- and post-test sessions. (C) Average threshold improvement in the high reward and low reward conditions. (D) Average slope at 75% accuracy across participants under the high reward and low reward conditions for the pre- and post-test sessions. (E) Average slope improvement in the high reward and low reward conditions. (F) Threshold index in the training session. (G) Correlation between the reward effects in the training and test sessions. (H) Accuracy from the reward awareness tests during pre- and post-test sessions. Error bars represent the standard error of the mean across participants.

Motion detection task: training

We analyzed the behavioural data during the training sessions. A repeated measure ANOVA showed no significant difference on threshold index (threshold index of training session i = threshold of session i / pre-test threshold) between the high reward and low reward conditions ($F(1,15) = 1.64, p = .22$, Figure 4F). However, when we further calculated the correlation between the reward-induced facilitation effects in the training and test sessions, we observed significant correlations. That is, the threshold index during the training (the averaged threshold index difference in two rewarding conditions across four training sessions) was significantly correlated with the reward-induced learning improvement in the test sessions ($r = 0.789, p < .001$, Figure 4G). These results suggest that reward played an important role in forming the learning effect, but such effect was not strongly reflected in the behavioural thresholds of the training sessions. One explanation of these results could be due to the influence of the reward feedback, i.e., the reward feedback existed in the training but not in the test sessions. Indeed, these results are in line with previous studies about the reward effect on visual attention (Anderson, Laurent, & Yantis, 2013), motor learning (Abe et al., 2011), and working memory (Gong & Li, 2014), which similarly reported that reward effect can be observed in the post-test session, but did not exist in the training phase when reward feedback was given. These findings suggest that the difference in reward delivery context between the training and test sessions may lead to the results that reward effect can be better observed during the test sessions after the training. However, the exact underlying mechanism is not clear yet and needs future investigations.

Reward awareness tests

The logic behind the reward awareness tests was that, if a reward was inaccessible to awareness and the existence of a reward-direction contingencies was not explicitly known by participants during the experiment, it would be reasonable to infer that there was no strategy-based attentional bias toward the high reward-associated stimuli during the training and post-test sessions.

Accuracy in both the pre-test reward awareness test (at the end of day 1) and the post-test reward awareness test (at the end of day 6) were not significantly different from chance level (mean = 0.51, SE = 0.009; $t_{16} = 1.53, p = .14$; mean = 0.49, SE = 0.01; $t_{16} = -1.00, p = .33$). The accuracy in the pre-test reward awareness test was not significantly different from that in the post-test ($t_{16} = 1.62, p = .12$, Figure 4H). All individual accuracies of the reward awareness tests were not significantly different from chance level (binomial test: all $p > .22$).

After informing the participants about the existence of reward-direction contingencies, the proportion of correctly guessing the contingency (12 out of 17) did not differ from chance level (binomial test: $p = .14$). The guessing accuracy seemed somewhat high though not significantly different from chance.

One possible reason was that some participants did not quite understand the question the first time and could not make a response within 3 s, so we had to give them a second trial. During the delay, participants had more than sufficient time to introspect their relative performance on the two motion directions in the post-training test, and thus increased the probability of correct guessing. Nevertheless, none of the participants reported that they had known the contingencies before we informed them, and most of the participants considered the rewards as being given randomly. Taken together, these results suggest that the reward value was unconscious to the participants during the training sessions, and the reward-direction contingencies were not explicitly known by the participants.

Further support for equivalent top-down attention for the two directions was the response time during training. The response time for the high reward direction did not differ significantly from that for the low reward direction either during the four days of training (repeated measured ANOVA showed main effect of training sessions, $F(3,45) = 5.87$, $p = .002$; but neither main effect of reward, $F(1,15) = 0.045$, $p = .83$; nor the interaction of reward and training day, $F(3,45) = 0.329$, $p = .80$), or in the pre-test and post-test sessions (repeated measured ANOVA showed main effect of stimulus signal-to-noise ratio, $F(4,64) = 75.06$, $p < .001$; main effect of test sessions, $F(1,16) = 19.79$, $p < .001$; but no main effect of reward, $F(1,16) = 0.151$, $p = .73$; nor interaction between reward and other two factors, all $p > .20$). In addition, the reward-facilitated learning effect could not be ascribed to different frequencies of reward delivery. The results revealed that the accuracy for the two motion directions did not differ in the four days of training (repeated measured ANOVA showed neither main effect of reward, $F(1,15) = 0.023$, $p = .88$; nor the interaction between reward and training session, $F(3,45) = 2.20$, $p = .10$), indicating equivalent number of trials for the delivery of high and low rewards. Finally, similar to the results in Experiment 1, participants rated 50 cents as making them happier than one cent during training (mean difference = 1.87, SE = 0.24; $t_{14} = 7.88$, $p < 10^{-5}$).

CONCLUSION

With two psychophysical experiments, our results demonstrate that an unconsciously presented monetary reward during the training can facilitate perceptual learning of coherent motion detection. Specifically, the results of Experiment 1 showed that a subliminally presented reward cue improved participants' sensitivity in the motion detection task. The results of Experiment 2 showed that using the same subliminal reward stimuli as feedback during four days of training, improvement of the sensitivity of motion detection was significantly larger for the high reward direction compared with the low reward direction.

We suggest that the results of Experiment 1 reflected the enhanced processing of the stimuli due to subliminal reward priming effect induced by masked coin images. This effect can be attributed to the initial activation of the reward-related neural circuits that led to unconscious elevation of motivation and recruitment of attentional resources (Bijleveld et al., 2009, 2011; Pessiglione et al., 2007). These results confirmed that the reward cue of masked coin image can be successfully processed unconsciously and validated its usage in Experiment 2.

The results from Experiment 2 suggest a facilitation effect of the reinforcement signal on perceptual learning. These findings agree with the previous perceptual learning models that suggested the important roles for reinforcement signal in gating the learning effect (Kahnt et al., 2011; Law & Gold, 2009; Roelfsema et al., 2010; Seitz & Watanabe, 2005). Specifically, under the framework of task-irrelevant perceptual learning, it was suggested that learning occurs as a result of the changes in the selectivity of sensory neurons that provided information to decision units, and this plasticity was enabled by the concurrently released widespread reinforcement signal. In our experiment, which can be categorized as task-relevant learning, correct responses on the training task were followed by the delivery of a subliminally presented monetary reward feedback. The difference between the predicted reward and the actually received reward defined a form of the reward prediction error signal (Schultz, 2007). Particularly, the predicted reward was equally uncertain for the two motion directions, as the reward-stimulus contingency could not be directly accessed by the participants. The reward prediction error was thus mainly determined by the magnitude of the actual reward the participants received after each response. If only the feedback on correctness was provided, learning on the two motion directions would be equivalent. However, introducing the subliminal high reward feedback to one condition resulted in a larger positive reward prediction error. Therefore, the connection between the sensory neurons and the decision units for this direction was better strengthened, leading to a more efficient weights updating for the readout process (Kahnt et al., 2011; Law & Gold, 2009). Together with previous studies, our results suggest that the task-relevant perceptual learning can occur as a result of a reinforcement learning process in which the reward prediction error is calculated based on the subliminal reward feedback. These findings further support the importance of the reward signal in both the task-relevant and task-irrelevant perceptual learning.

Importantly, with a reward learning paradigm, our study differed from the previous investigations in two critical ways. First, a reward feedback was used instead of neuromodulator-elevated pharmacological treatment, in which the reinforcement signal is not a phasic signal and may mediate learning through the sustained enhancement of visual attention. Second, we used the unconsciously presented monetary images as an exogenous reward feedback. It is well known that consciously perceived exogenous motivational incentives can enhance attentional allocation and improve learning. Previous studies have also shown

that endogenous reward, such as the sense of achievement after correct performance in one task, can concurrently act on the early visual processing and learning of task-irrelevant stimulus (Pascucci & Turatto, 2013; Seitz & Watanabe, 2003; Seitz et al., 2005). Our results demonstrate an alternative approach in which the reinforcement signal evoked by unconscious monetary reward feedback could also improve the perceptual learning. However, we suggest that the sense of endogenous achievement during the training may also play roles in the observed learning effects in our experiments and other studies with explicit feedback on task performance. Third, we used an unconscious, rather than a conscious, reward to minimize the possible strategy-based attentional bias and transient attentional enhancement on the visual stimuli. As revealed by the reward awareness tests, the participants were unaware of the reward value throughout the training sessions and did not realize the existence of the reward-direction contingencies. As the reward feedback was presented after the stimuli and response, the transient enhancement of attentional bias was also minimized.

Moreover, the 2AFC task was used instead of choosing between directions of the two tested stimuli. Therefore, the facilitation in the high reward direction reflected sensitivity changes but not the artefact of choice bias due to stimulus-response association. However, it must be noted that an unconsciously presented reward induces a weaker reinforcement signal than a conscious reward. We used the supraliminal reward cueing practice in Experiment 1 and the oddball detection task in Experiment 2 as controls to maximize the subliminal reward processing. Nevertheless, we could not completely exclude the possible contribution from the subliminal motivation delivered between trials, though the experimental procedure had largely blocked the conscious motivation such as deliberate and strategic attentional employment. Future studies with novel procedures are required to quantitatively investigate the contribution of the subliminal motivation in perceptual learning.

In previous perceptual learning studies, it was well documented that concurrent improvement on multiple stimuli occurred, when the training procedure were an ABAB design or the stimuli were shown in a fixed sequence (Yu, Klein, & Levi, 2004; Zhang et al., 2008). There was no reward feedback in these studies. In our Experiment 2, the low reward condition showed a trend of significance on threshold decrease but no significant effect on slope increase. These results suggest that there was a little effect of learning, if any, for the low reward condition. In fact, the low reward we offered in the experiment can be neglected as a non-reward, as one Chinese cent has very low value to nowadays market and has been taken out of circulation in people's daily life. The results of the happiness rating showed an approach-to-zero happiness level for the low reward. In this regard, the observed results for the low reward condition are consistent with previous studies in visual perceptual learning and motor learning, which showed an improvement in learning only in the rewarded but not the non-reward condition (Abe et al., 2011; Seitz et al.,

2009). Inhibition mechanism may play an important role in this phenomenon, but further investigations are required to clarify this issue.

In conclusion, the results suggest that trial-by-trial unconscious feedback of a monetary reward during training improves the motion perceptual learning. These results provide behavioural evidence that the reward signal directly modulates task-relevant perceptual learning. Combined with previous studies, our findings suggest that reward signal plays an important modulatory role in both task-relevant and task-irrelevant perceptual learning.

Funding

This work was supported by the National Natural Science Foundation of China [No. 31271081, 31230029], the National High Technology Research and Development Program of China (863 Program) [No. 2012AA011602], the Program for New Century Excellent Talents in University, State Education Ministry.

REFERENCES

- Abe, M., Schambra, H., Wassermann, E. M., Luckenbaugh, D., Schweighofer, N., & Cohen, L. G. (2011). Reward improves long-term retention of a motor memory through induction of offline memory gains. *Current Biology*, *21*(7), 557–562. doi:10.1016/j.cub.2011.02.030
- Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(12), 5718–5722. doi:10.1073/pnas.90.12.5718
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(25), 10367–10371. doi:10.1073/pnas.1104047108
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2013). Reward predictions bias attentional selection. *Frontiers in Human Neuroscience*, *7*, 262. doi:10.3389/fnhum.2013.00262
- Bijleveld, E., Custers, R., & Aarts, H. (2009). The unconscious eye-opener: Pupil size reveals strategic recruitment of resources upon subliminal reward cues. *Psychological Science*, *20*(11), 1313–1315. doi:10.1111/j.1467-9280.2009.02443.x
- Bijleveld, E., Custers, R., & Aarts, H. (2010). Unconscious reward cues increase invested effort, but do not change speed-accuracy tradeoffs. *Cognition*, *115*(2), 330–335. doi:10.1016/j.cognition.2009.12.012
- Bijleveld, E., Custers, R., & Aarts, H. (2011). Once the money is in sight: Distinctive effects of conscious and unconscious rewards on task performance. *Journal of Experimental Social Psychology*, *47*(4), 865–869. doi:10.1016/j.jesp.2011.03.002
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211. doi:10.1016/j.tics.2006.03.007
- Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science*, *17*(3), 222–227. doi:10.1111/j.1467-9280.2006.01689.x
- Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science*, *20*(6), 778–784. doi:10.1111/j.1467-9280.2009.02360.x

- Dosher, B. A., & Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(23), 13988–13993. doi:10.1073/pnas.95.23.13988
- Dosher, B. A., & Lu, Z. L. (1999). Mechanisms of perceptual learning. *Vision Research*, *38*(9), 3197–3221. doi:10.1016/S0042-6989(97)00273-3
- Frankó, E., Seitz, A. R., & Vogels, R. (2009). Dissociable neural effects of long term stimulus-reward pairing in macaque visual cortex. *Journal of Cognitive Neuroscience*, *22*(7), 1425–1439. doi:10.1162/jocn.2009.21288
- Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, *31*(5), 681–697. doi:10.1016/S0896-6273(01)00424-X
- Gong, M., & Li, S. (2014). Learned reward association improves visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 841–856. doi:10.1037/a0035131
- Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience*, *30*(33), 11096–11103. doi:10.1523/JNEUROSCI.1026-10.2010
- Hickey, C., Chelazzi, L., & Theeuwes, J. (2011). Reward has a residual impact on target selection in visual search, but not on the suppression of distractors. *Visual Cognition*, *19*(1), 117–128. doi:10.1080/13506285.2010.503946
- Kahnt, T., Grueschow, M., Speck, O., & Haynes, J. D. (2011). Perceptual learning and decision-making in human medial frontal cortex. *Neuron*, *70*(3), 549–559. doi:10.1016/j.neuron.2011.02.054
- Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(11), 4966–4970. doi:10.1073/pnas.88.11.4966
- Law, C. T., & Gold, J. I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature Neuroscience*, *12*(5), 655–663. doi:10.1038/nn.2304
- Maunsell, J. H. (2004). Neuronal representations of cognitive state: Reward or attention? *Trends in Cognitive Sciences*, *8*(6), 261–265. doi:10.1016/j.tics.2004.04.003
- Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science*, *13*(5), 416–424. doi:10.1111/1467-9280.00474
- Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, *8*(6), 2201–2211.
- O'Brien, J. L., & Raymond, J. E. (2012). Learned predictiveness speeds visual processing. *Psychological Science*, *23*(4), 359–363. doi:10.1177/0956797611429800
- Pascucci, D., & Turatto, M. (2013). Immediate effect of internal reward on visual adaptation. *Psychological Science*, *24*(7), 1317–1322. doi:10.1177/0956797612469211
- Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., & Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron*, *59*(4), 561–567. doi:10.1016/j.neuron.2008.07.005
- Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the brain translates money into force: A neuroimaging study of subliminal motivation. *Science*, *316*(5826), 904–906. doi:10.1126/science.1140459
- Raio, C. M., Carmel, D., Carrasco, M., & Phelps, E. A. (2012). Nonconscious fear is quickly acquired but swiftly forgotten. *Current Biology*, *22*(12), R477–479. doi:10.1016/j.cub.2012.04.023
- Roelfsema, P. R., van Ooyen, A., & Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences*, *14*(2), 64–71. doi:10.1016/j.tics.2009.11.005
- Sagi, D., & Tanne, D. (1994). Perceptual learning: Learning to see. *Current Opinion in Neurobiology*, *4*(2), 195–199. doi:10.1016/0959-4388(94)90072-8

- Sasaki, Y., Nanez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience*, *11*(1), 53–60. doi:10.1038/nrn2737
- Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature*, *412*, 549–553. doi:10.1038/35087601
- Schultz, W. (2007). Behavioral dopamine signals. *Trends in Neuroscience*, *30*(5), 203–210. doi:10.1016/j.tins.2007.03.007
- Seitz, A. R., & Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? *Nature*, *422*, 36. doi:10.1038/422036a
- Seitz, A. R., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences*, *9*(7), 329–334. doi:10.1016/j.tics.2005.05.010
- Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, *61*(5), 700–707. doi:10.1016/j.neuron.2009.01.016
- Seitz, A. R., Lefebvre, C., Watanabe, T., & Jolicoeur, P. (2005). Requirement for high-level processing in subliminal learning. *Current Biology*, *15*(18), R753–755. doi:10.1016/j.cub.2005.09.009
- Watanabe, M., Cheng, K., Murayama, Y., Ueno, K., Asamizuya, T., Tanaka, K., & Logothetis, N. (2011). Attention but not awareness modulates the BOLD signal in the human V1 during binocular suppression. *Science*, *334*(6057), 829–831. doi:10.1126/science.1203161
- Watson, A. B., & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*(2), 113–120. doi:10.3758/BF03202828
- Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*(8), 1314–1329. doi:10.3758/BF03194545
- Xiao, L. Q., Zhang, J. Y., Wang, R., Klein, S. A., Levi, D. M., & Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology*, *18*(24), 1922–1926. doi:10.1016/j.cub.2008.10.030
- Yu, C., Klein, S. A., & Levi, D. M. (2004). Perceptual learning in contrast discrimination and the (minimal) role of context. *Journal of Vision*, *4*(3), 169–182. doi:10.1167/4.3.4
- Zedelius, C. M., Veling, H., & Aarts, H. (2011). Boosting or choking—how conscious and unconscious reward processing modulate the active maintenance of goal-relevant information. *Consciousness and Cognition*, *20*(2), 355–362. doi:10.1016/j.concog.2010.05.001
- Zhang, J. Y., Kuai, S. G., Xiao, L. Q., Klein, S. A., Levi, D. M., & Yu, C. (2008). Stimulus coding rules for perceptual learning. *PLoS Biology*, *6*(8), e197. doi:10.1371/journal.pbio.0060197
- Zhang, J. Y., Zhang, G. L., Xiao, L. Q., Klein, S. A., Levi, D. M., & Yu, C. (2010). Rule-based learning explains visual perceptual learning and its specificity and transfer. *Journal of Neuroscience*, *30*(37), 12323–12328. doi:10.1523/JNEUROSCI.0704-10.2010