Visual Cognition, 2015
Vol. 23, No. 1–2, 161–178, http://dx.doi.org/10.1080/13506285.2014.981625

Routledge
Taylor & Francis Group

Unconscious reward facilitates motion perceptual learning

Xin Xue[1,2,3], Xiaolin Zhou[1,2,3,4], and Sheng Li[1,2,3]

[1]Department of Psychology, Peking University, Beijing, China
[2]Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China
[3]PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China
[4]Center for Brain and Cognitive Sciences, Peking University, Beijing, China

(*Received 1 June 2014; accepted 23 October 2014*)

Reward signals play an important role in guiding human learning behaviour. Recent studies have provided evidence that reward signals modulate 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 signals play 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.

Downloaded by [Peking University] at 17:53 18 June 2015

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 literature

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

© 2015 Taylor & Francis

Downloaded by [Peking University] at 17:53 18 June 2015

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 (Dosher & 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 results 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 stimuli 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.

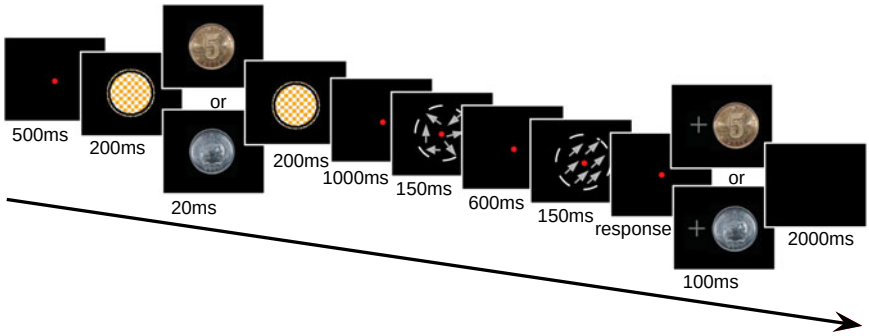

Figure 1. Schematic illustration of the procedure for Experiment 1. For each trial, participants received the monetary reward indicated by the coin image masked by orange-and-white checkerboards if a correct response was made for the motion detection task.

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

Downloaded by [Peking University] at 17:53 18 June 2015

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

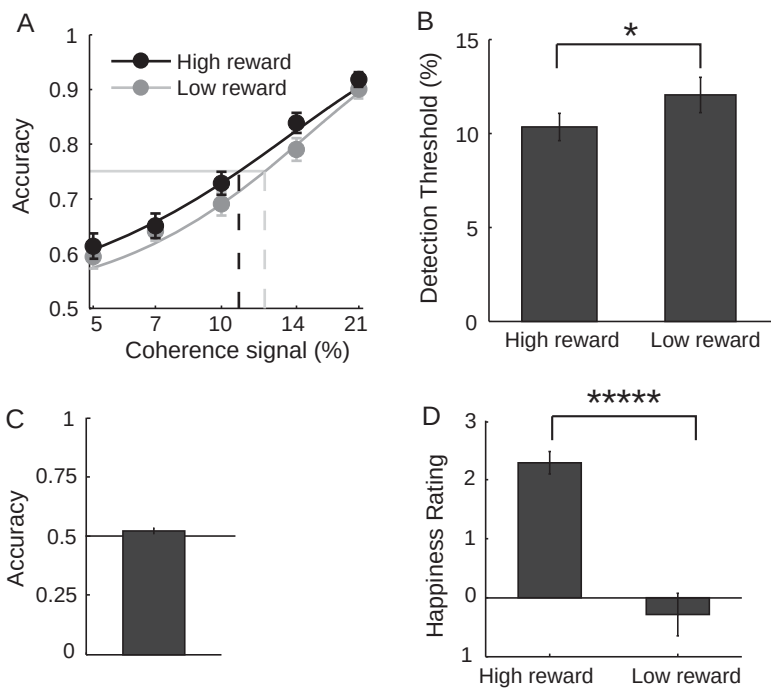

Figure 2. Results of Experiment 1. (A) Psychometric functions fitted for the average accuracy across all participants in the high-reward (black) and low-reward (grey) conditions. (B) Average detection threshold at 75% accuracy across participants under high-reward and low-reward conditions. (C) Chance-level 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.

Downloaded by [Peking University] at 17:53 18 June 2015

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 consciously 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.

Downloaded by [Peking University] at 17:53 18 June 2015

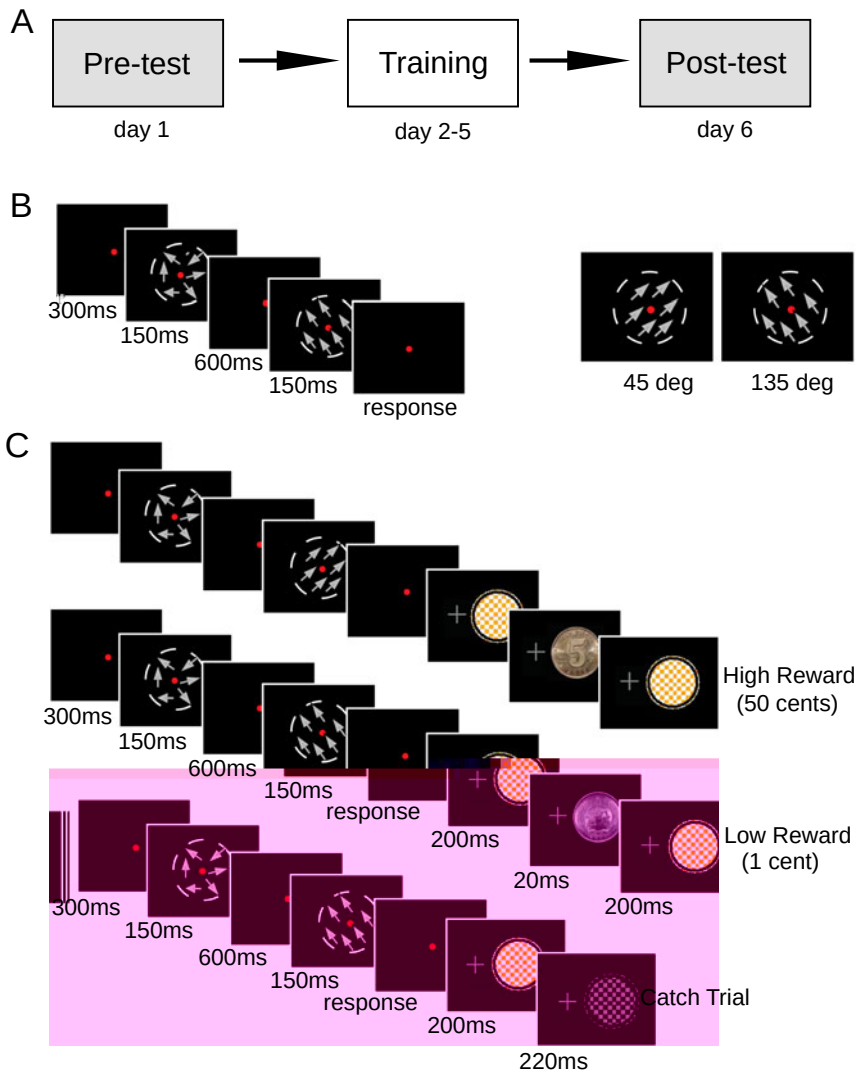

Figure 3. Schematic illustration of the procedure for Experiment 2. (A) General procedure. (B) Procedure of a trial in the pre-test and post-test sessions (left). Dynamic random dots of two directions (45° and 135°) were used in the experiment (right). (C) Procedures of three trial types in the training 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

Downloaded by [Peking University] at 17:53 18 June 2015

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 tasks 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 responses.

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

Downloaded by [Peking University] at 17:53 18 June 2015