UC Irvine
UC Irvine Previously Published Works

Title

High reward enhances perceptual learning

Permalink

Journal

ISSN

Authors

Zhang, Pan
Hou, Fang
Yan, Fang-Fang
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

High reward enhances perceptual learning

Pan Zhang

CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
Department of Psychology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China
Laboratory of Brain Processes (LOBES), Center for Cognitive and Brain Sciences, Center for Cognitive and Behavioral Brain Imaging, and Departments of Psychology, The Ohio State University, Columbus, OH, USA

Fang Hou

School of Ophthalmology & Optometry and Eye Hospital, Wenzhou Medical University, Wenzhou, Zhejiang, China

Fang-Fang Yan

CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
Department of Psychology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

Jie Xi

CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
Department of Psychology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

Bo-Rong Lin

CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
Department of Psychology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

Jin Zhao

CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
Department of Psychology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

Jia Yang

CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
Department of Psychology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

Ge Chen

CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
Department of Psychology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China
School of Arts and Design, Zhengzhou University of Light Industry, Zhengzhou, Henan, China

Meng-Yuan Zhang   CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

Qing He   CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
Department of Psychology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

Barbara Anne Dosher   Department of Cognitive Sciences and Institute of Mathematical Behavioral Sciences, University of California, Irvine, CA, USA ✉

Zhong-Lin Lu   Laboratory of Brain Processes (LOBES), Center for Cognitive and Brain Sciences, Center for Cognitive and Behavioral Brain Imaging, and Departments of Psychology, The Ohio State University, Columbus, OH, USA ✉

Chang-Bing Huang   CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
Department of Psychology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China ✉

Studies of perceptual learning have revealed a great deal of plasticity in adult humans. In this study, we systematically investigated the effects and mechanisms of several forms (trial-by-trial, block, and session rewards) and levels (no, low, high, subliminal) of monetary reward on the rate, magnitude, and generalizability of perceptual learning. We found that high monetary reward can greatly promote the rate and boost the magnitude of learning and enhance performance in untrained spatial frequencies and eye without changing interocular, interlocation, and interdirection transfer indices. High reward per se made unique contributions to the enhanced learning through improved internal noise reduction. Furthermore, the effects of high reward on perceptual learning occurred in a range of perceptual tasks. The results may have major implications for the understanding of the nature of the learning rule in perceptual learning and for the use of reward to enhance perceptual learning in practical applications.

Introduction

Perceptual expertise is critical for the survival of animals and humans. Acquiring perceptual expertise is usually time-consuming and often specific to the trained tasks and settings (Fahle & Poggio, 2002; Goldstone, 1998; Lu, Hua, Huang, Zhou, & Dosher, 2011; Sagi, 2011). How to learn more, learn more quickly, and generalize more broadly to untrained conditions is an unsettled question. Studies of perceptual learning have revealed important plasticity in adult humans that has become an integral component of our understanding of perception (Petrov, Dosher, & Lu, 2005; Sagi, 2011; Sasaki, Náñez, & Watanabe, 2012) and have led to noninvasive rehabilitation methods for a variety of clinical conditions (Lu et al., 2011; Sagi, 2011; Sasaki, Nanez, & Watanabe, 2010; Skinner, 1938). Previous studies have shown that more training increases the magnitude of learning but also specificity to the trained task, while cross-training on multiple tasks ("double training") can improve generalization across retinal locations (Jeter, Dosher, Liu, & Lu, 2010; Xiao et al., 2008). Here we used a novel compound monetary reward structure to investigate the effects of different types and magnitudes of reward on the rate of perceptual learning in trained tasks and generalizability across stimuli and eye in a range of simple visual tasks. To our knowledge, this is the first investigation that shows titrated effects of reward magnitude in perceptual learning in addition to informational feedback.

Reward plays a central role in incentive-based learning and development of goal-directed behaviors, engaging similar reward networks in humans and primates (Arias-Carrión & Pöppel, 2007; Dayan & Balleine, 2002; Haber & Knutson, 2010; Maunsell, 2004). Reward (or punishment) has been used to

improve task performance and/or learning in both animals and humans (Della Libera & Chelazzi, 2006; Kennerley & Wallis, 2009; Kim et al., 2015; Seitz, Kim, & Watanabe, 2009; Watanabe, Náñez, & Sasaki, 2001; Xue, Zhou, & Li, 2015).

The purpose of the current study is to investigate the differential impacts of different reward manipulations, independent of the information value of trial-by-trial feedback. We also aim to discover any differential consequences of these reward manipulations on generalizability, and how it relates to the magnitude of learning in the trained perceptual task.

There is extensive literature on the complex role of different forms of feedback—information about response accuracy in the task—on visual perceptual learning and a somewhat smaller literature on the role of reward. Although trial-by-trial feedback has been the norm in visual perceptual learning, training improvements can occur with block feedback or even without feedback (or reward) in a range of task conditions (e.g., Herzog & Fahle, 1997; Liu, Lu, & Dosher, 2010, 2012; Petrov et al., 2005; Petrov, Dosher, & Lu, 2006; Dosher, Jeter, Liu, & Lu, 2013; see Liu, Dosher, & Lu, 2014, 2015, Lu & Dosher, 2009 and Dosher & Lu, 2017; for detailed reviews).

There are previous studies investigating reward in visual perceptual learning. One set of studies examined task-irrelevant learning, in which properties of a near-subliminal visual stimulus are learned when it is temporally paired with a target in a primary task (e.g., Kim et al., 2015; Seitz et al., 2009). In these tasks, perceptual learning is seen as operating through a diffused endogenous reward signal associated with the success in the primary task (Della Libera & Chelazzi, 2006; Pascucci & Turatto, 2013; Seitz et al., 2009; Seitz et al., 2005; Watanabe et al., 2001). Perceptual learning has also been documented in response to rewards only (e.g., juice drops) that substitute for other informational feedback (Seitz et al., 2009). Subliminal reward has been shown to have a modest effect on the magnitude of perceptual learning in motion direction discrimination (Xue et al., 2015).

These studies, especially those that used either primary or secondary rewards (e.g., juice for thirsty participants or symbols of monetary reward, respectively), indicate that perceptual learning can occur in response to reward signals. Yet questions remain: Does reward operate differently from simple feedback? Does reward magnitude influence learning and generalization? Does subliminal reward operate differently? Understanding the impact of different types and magnitudes of reward on the speed, magnitude, and/or generalization of perceptual learning is not only important for our understanding of the nature of visual plasticity, especially the learning rule (Dayan & Balleine, 2002), but also for practical applications of

perceptual learning. Although the earlier literature emphasized specificity of visual perceptual learning, an increasing number of recent papers have focused on generalization of perceptual learning (Sasaki et al., 2012), which is important for the development of expertise or remediation of visual conditions. Optimizing perceptual learning, including the rate and magnitude of learning and the degree of generalization, may involve development of better reward structures (Dosher & Lu, 2017).

In the current study, we developed a new compound reward structure that consisted of combinations of trial-by-trial, between-block, between-session, and absolute performance rewards. Together, these rewards are translated using a conversion rate from points to monetary compensation (Figure 1A and B), all secondary rewards. We evaluated the role of reward as distinct from informational feedback about accuracy, which is provided in all conditions, in four experiments.

Experiment 1 examined the differential effectiveness of visual perceptual learning in five reward conditions: no reward, block reward, low reward, high reward, and subliminal reward in a contrast-detection task. The no-reward condition includes no reward points, and compensation is equal to base pay for participation. The block condition delivers currency images (Chinese yuan) and text rewards for improvements in performance between blocks and sessions or for achieving absolute performance criteria. The high, low, and subliminal conditions delivered trial-by-trial reward signals in addition to block and session rewards. In the subliminal condition, currency images are brief and of low contrast, and the block and session rewards are supraliminal. In the high, subliminal, and block conditions, base pay is low, and the conversion from reward points to payoff is high; in the low-reward condition, base pay is the same as the no-reward condition, and the conversion from reward points to payoff is low. Base and performance-related translation from reward to compensation was set to approximately equate pay across conditions in an effort to equate overall motivation while, at the same time, amplifying the consequences of performance achievement (see Methods). These reward conditions are designed to span a range of effectiveness with the no-reward condition serving as a baseline and the high-reward condition including the strongest performance-related rewards in the experiment. This also permits paired condition comparisons that isolate the effectiveness of different reward features. Experiments 2–4 focus on comparing high reward to no reward to investigate other questions. Experiment 2 focuses on learned improvements under different external noise titrations in contrast detection to identify the mechanism of the learned improvements. Experiments 3 and 4 examine learned improvements in Vernier hyperacuity and
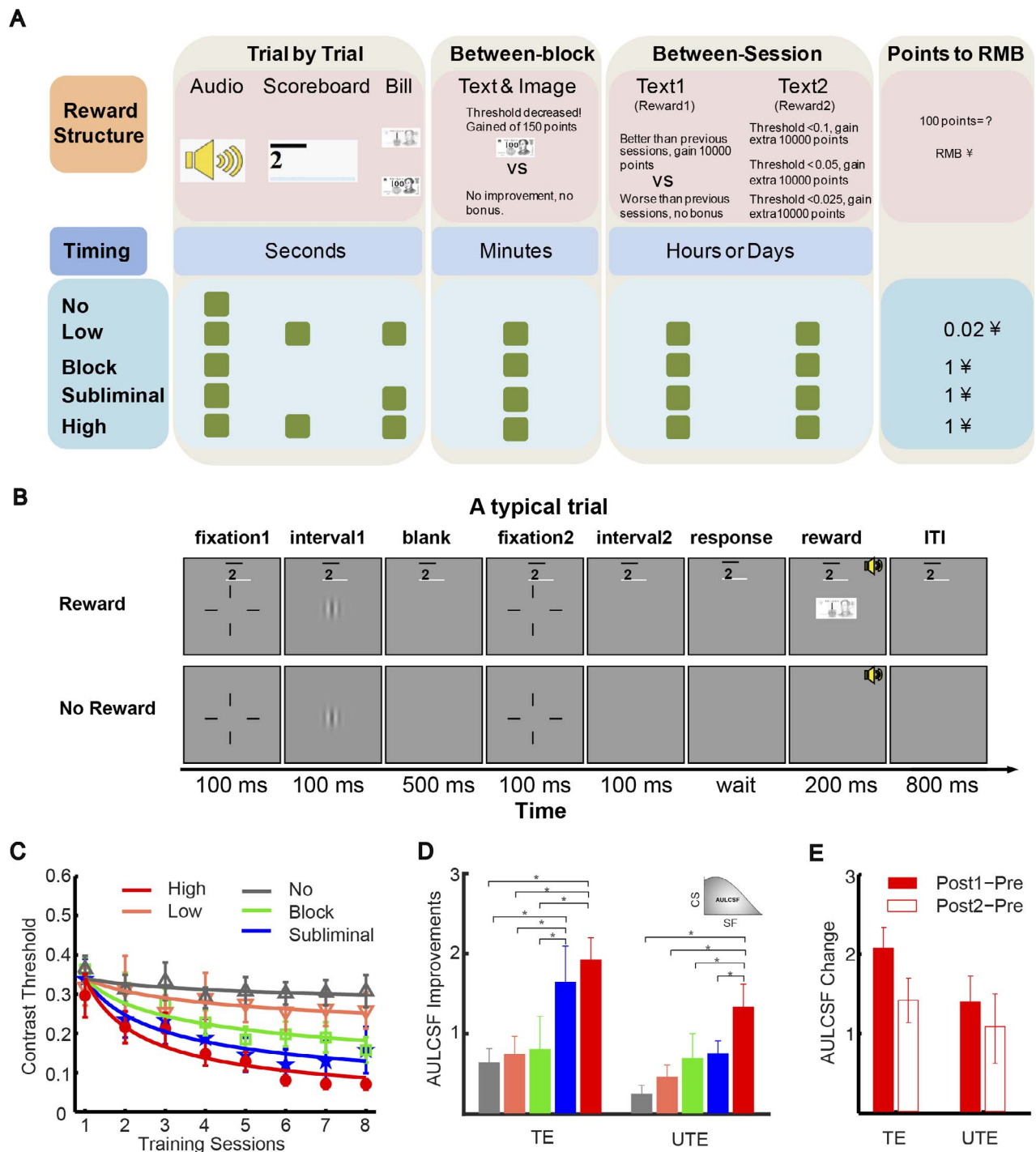
Figure 1. (A) Reward structure. Five reward groups (no, low, block, subliminal, and high reward) differed in the combination of rewards in multiple time scales; the visibility of the trial-by-trial reward signal (seconds), block reward signals (minutes), or session reward signals (hours or day); and/or the conversion rate of reward points to monetary compensation (see Methods for details). (B) Training procedure. A two-interval, forced-choice procedure was used for training and assessment of contrast sensitivity (see Methods). (C) Learning curves—contrast threshold as a function of training session—were fit by power functions with different learning rates for the five reward groups. (D) Schematic diagram of the area under log contrast sensitivity function (AULCSF) and the improvement of AULCSF in trained (TE) and untrained eyes (UTE). Together with panel C, gray, light red, green, blue, and red bars and curves denote data in no-, low-, block-, subliminal-, and high-reward conditions, respectively. CS: contrast sensitivity; SF: spatial frequency. (E) Retention of improved AULCSF. Solid and hollow red bars denote AULCSF improvements in the first and second posttests, respectively. Error bars are the standard error.

global motion-detection discrimination, respectively (with block and session rewards slightly modified to suit the dependent measures in those tasks). Each experiment also measures generalization to other tasks or conditions: pretraining and posttraining measures of monocular contrast sensitivity function in the trained and untrained eyes (Experiments 1 and 2), of the trained and untrained locations (Experiment 3), and in the trained and untrained motion directions (Experiment 4).

We found that high monetary reward can greatly boost the rate and the magnitude of perceptual learning. Generalization of perceptual learning of contrast detection from a trained spatial frequency to other spatial frequencies increased with the magnitude of learning as did interocular, interlocation, or interdirection transfer. Perceptual learning with high reward enhanced internal noise reduction. Finally, the benefits of high reward occurred in a range of perceptual tasks. The results may have major implications for the understanding of the nature of the learning rule in perceptual learning and for the use of reward to enhance perceptual learning in practical applications.

Experiment 1: Monetary reward modulates contrast detection learning and generalization

Experiment 1 examines the relative effectiveness of five different conditions of monetary reward—no reward, block reward, low reward, high reward, and subliminal reward as described above (see also Methods)—on the rate and magnitude of learning contrast detection at a trained spatial frequency in the nondominant eye. To assess generalization of training, we also measured the thresholds in contrast detection of a sine-wave patch at various spatial frequencies, which defines the contrast sensitivity function (CSF) before and after training and in both the trained and untrained eyes. All observers received auditory feedback on response accuracy in each trial, so the titrated impact of reward on learning and generalization is separated from the effects of information.

Methods

Observers

Forty-one observers (23.12 ± 0.29 years) with normal or corrected-to-normal vision and informed written consent participated in this study. None were aware of the purpose of the study. The work was carried out in accordance with the Declaration of Helsinki.

Apparatus

The study was conducted on a PC computer running MATLAB programs with PsychToolbox extensions (Brainard, 1997; Pelli, 1997). Stimuli were displayed on a Sony G220 monitor with a $1,600 \times 1,200$ pixel resolution, 85 Hz frame rate, and 36 cd/m^2 background luminance. A special circuit was used to combine two eight-bit output channels of the video card to produce 14 bits of gray levels (X. Li, Lu, Xu, Jin, & Zhou, 2003). Observers placed their head on a chin rest and viewed the displays monocularly with an opaque patch on the other eye. The display subtended $8.33° \times 6.25°$ at a viewing distance of 2.76 m.

Design

The experiment consisted of three phases: (a) pretraining measurements of the monocular CSF in both eyes (two sessions of seven blocks of 100 trials), (b) training in the sine-wave grating detection task at each individual observer's cutoff spatial frequency (estimated from pretraining measurements) in the nondominant eye (eight daily sessions, each with seven blocks of 80 trials), and (c) posttraining measurements of the monocular CSF in both eyes (two sessions, each containing seven blocks of 100 trials). For each monocular CSF, contrast detection was assessed at spatial frequencies of 1, 2, 4, 8, 16, 24, and 32 c/° with seven interleaved three-down/one-up staircases that converge to 79.4% correct using 100 trials per staircase (see Procedure). Stimuli were vertical sinusoidal gratings of 2° in diameter with a half-Gaussian ramp ($\sigma = 0.25°$). Observers ran about 100 practice trials before data collection.

The compound reward structure (see Figure 1A) could include trial-by-trial reward messages (a scoreboard and image of a Chinese currency bill), between-block reward messages (text plus a currency bill image if the block performance had improved), and between-session rewards (text indicating different reward points for performance better than the last session, than all sessions, or above absolute performance criteria). The five reward conditions, as described above, invoked different combinations or types of reward: (a) *high reward* ($n = 8$) with all reward signals and a high conversion rate, (b) *low reward* ($n = 7$) with all reward signals and a low conversion rate, (c) *no reward* ($n = 11$) with no reward signals and only base pay, (d) *block reward* ($n = 9$) with block reward signals and a high conversion rate, and (e) *subliminal reward* ($n = 6$) in which the trial-by-trial reward signal was subliminal (low contrast and briefly presented currency image and no scoreboard) but was otherwise equivalent to the high-reward condition. The reward conditions were manipulated between groups. (The sample sizes reflect random assignment except that, in the no-reward

condition, we targeted a slightly higher sample size to capture a low rate of learning.)

Observers received base pay for participating in the experiment. In addition, observers except those in the no-reward condition gained reward points for performance in trials, blocks, and sessions during the training phase. Total payment was the sum of the base pay and performance-dependent reward. The base pay was 35¥ (Chinese yuan), 35¥, 10¥, 10¥, and 10¥ in the no-, low-, block-, subliminal-, and high-reward conditions, respectively. In trial-by-trial rewards, each correct response was awarded two reward points with four extra reward points following three consecutive correct responses. In block rewards, observers were awarded 150 reward points if the threshold of the current block was lower than that of previous blocks in the same session. In session rewards, observers gained 400 reward points if the final threshold of the session was lower than that of the previous session or 1,000 reward points if it was lower than that of all previous sessions (session reward 1, Figure 1A). They also received an extra 10,000, 20,000, and 30,000 reward points if the threshold in a session was lower than 0.1, 0.05, and 0.025 (in contrast unit), respectively (session reward 2, Figure 1A).

Observers were instructed to gain as many reward points as they could. In the high-, block-, and subliminal-reward conditions, observers redeemed 100 reward points for 1¥ (exchange rate = 100:1). In the low-reward condition, observers redeemed 100 reward points for only 0.02¥ (exchange rate = 5,000:1). That is, the block-, subliminal-, and high-reward groups experienced the same conversion rate (or magnitude of reward), which is higher than that of the low-reward group. Observers were informed of the relevant compound rewards and conversion rule before the first training session. For subjects in the subliminal group, the existence of invisible trial-by-trial reward was not disclosed, and they were informed only of the block and session rewards and the conversion formula; that is, observers in the block and subliminal groups received the same instructions.

Images of 1¥ and 100¥ bills were matched in size ($1.30° \times 0.78°$) and contrast. A 4.4% root mean square (RMS) contrast was used in the subliminal-reward condition; a 13.7% RMS contrast was used in all other conditions. For the subliminal condition, the following tests were performed: (a) Before and after CSF tests, observers were told that 10 pictures of uniform luminance will be displayed in sequence, and they were asked to check display quality by eye at half the viewing distance used in the main experiment; the bill image used in the subliminal-reward condition was presented in the center of the display. All observers reported that the display was homogenous in luminance, and no one ever reported they detected any other visual patterns.

(b) Observers were also required to verbally report to the experimenter if they saw anything other than the fixation and sine-wave grating during either test or training sessions. No observer reported any pattern(s) during the experiment.

All observers filled out the sensitivity to reward and punishment questionnaire (SRPQ), consisting of 48 simple questions related to self-evaluations of reward and punishment responses (Supplementary Table S1). Finally, we evaluated the retention of contrast-sensitivity performance in each eye for seven of the eight observers in the high-reward condition at least five months after the training and assessment sessions.

Procedure

A two-interval, forced-choice sine-wave grating detection task was used for training and for assessment of the CSF (Figure 1B). Thresholds at 79.4% correct were measured using a three-down/one-up staircase procedure (Levitt, 1971). A target grating was presented with equal probability in one of two 100-ms temporal intervals, separated by 500 ms. Observers were asked to indicate which interval contained the grating by pressing a key on the computer keyboard. No feedback was provided in the pretraining and posttraining assessments of CSF. In the training task in all reward conditions, feedback consisted of an auditory beep after each correct response. This equates the information about accuracy of the response in all reward conditions. In the high and low trial-by-trial reward conditions, a correct response was followed by a currency image, a black bar, and a number in which each increment of the black bar represented one reward point and the number represented multiples of 100 points. A comparison white bar, denoting 100 points, was also shown. A high-contrast ¥1 picture was shown for 200 ms following a correct response, or a high-contrast ¥100 picture was shown after three consecutive correct responses. In the block-reward condition, no currency image was shown after each trial; in the subliminal-reward condition, a low-contrast ¥1 or ¥100 picture was shown for 16 ms after each correct response. In the subliminal and block conditions, the length of the black bar increased randomly, and the observers were so informed.

Data analysis

Unless noted otherwise, the significance level was $p < 0.05$, and marginal significance corresponded to $0.05 \leq p < 0.10$ throughout the paper.

A power function, $C(t) = C_0 \, t^{-\rho}$, was used to fit the average learning curves, where C_0 is the initial contrast threshold, t is the training session number, and ρ is the learning rate (Dosher & Lu, 2005). A nonlinear least-

square method, implemented in MATLAB (Math-Works, Natick, MA), was used to minimize the sum of squared differences between model predictions and observed values, $SSE = \sum (y_i - \hat{y}_i)^2$, and the goodness of fit was gauged by r^2:

$$r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}, \quad (1)$$

where y_i and $\hat{y}_i$ represent the observed and predicted values, respectively, and $\bar{y}$ is the mean of all the observed values.

We compared the learning curves in the five reward conditions in a nested-model testing framework. The model lattice consisted of 11 models, ranging from the full 10-parameter model with independent C_0 and ρ in the five conditions to the most reduced two-parameter model with identical C_0 and ρ across all five (see Supplementary Table S1 for details).

An F test was used to statistically compare the fits of nested models:

$$F(df_1, df_2) = \frac{\left(r^2_{full} - r^2_{reduced}\right)/df_1}{\left(1 - r^2_{full}\right)/df_2}, \quad (2)$$

where $df_1 = k_{full} - k_{reduced}$, $df_2 = N - k_{full}$, k_{full} and $k_{reduced}$ are the numbers of parameters of the full and reduced models, respectively, and N is the number of data points. The model that was statistically equivalent to the full model and superior to all its reduced forms was defined as the best-fitting model.

The log contrast sensitivity function graphs log contrast sensitivity (1/threshold) as a function of spatial frequency (see inset Figure 1D). The area under the log contrast sensitivity function (AULCSF), which provides a broad measure of contrast sensitivity across all spatial frequencies, was calculated to evaluate the improvement in each eye (Koop, Applegate, & Howland, 1996).

Results

Sensitivity to reward and punishment

An analysis of variance (ANOVA) performed on the responses to the SRPQ indicated that observers in the five reward conditions were comparable in the tendency to seek reward, $F(4, 36) = 1.435$, $p = 0.242$, and avoid punishment, $F(4, 36) = 0.663$, $p = 0.662$. There were no notable patterns relating the questionnaire responses and performance or learning in the behavioral task (all $ps > 0.1$).

Learning rate

Figure 1C shows the contrast threshold learning curves for the five reward conditions: no, low, block,

subliminal, and high reward, together with the best-fitting power function learning curves. The learning model that has different learning rates (ρ) but the same initial threshold C_0 in the five reward conditions provided the best fit ($1C_0$, 5ρ; Supplementary Table S1). This six-parameter model is statistically equivalent to the full model ($5C_0$, 5ρ; $r^2 = 92.97\%$ vs. $r^2 = 94.27\%$), $F(4, 30) = 1.704$, $p = 0.175$, and provided significantly better fits than more reduced forms (all $ps < 0.05$ in nested model tests; Supplementary Table S1). The high trial-by-trial reward condition produced the highest learning rate, followed by the subliminal trial-by-trial reward condition, the block-reward condition, and then the low trial-by-trial reward and no-reward conditions. The corresponding learning rates (mean $\pm$ SE) were 0.74 ± 0.01, 0.45 ± 0.02, 0.31 ± 0.01, 0.15 ± 0.01, and 0.09 ± 0.01, respectively. The standard errors were estimated using a bootstrap method with 1,000 iterations.

CSF improvements

Training at the cutoff spatial frequency improved contrast sensitivity over a range of spatial frequencies (Huang, Zhou, & Lu, 2008). A one-way ANOVA was performed on the increase of AULCSF in the trained eye with reward condition (no, low, block, subliminal, and high) as a between-subjects factor. Training in the no-, low-, block-, subliminal-, and high-reward conditions increased the AULCSF in the trained eye by 0.64 ± 0.18, 0.74 ± 0.22, 0.80 ± 0.41, 1.64 ± 0.45, and 1.92 ± 0.28 $\log_{10}$ units, respectively (Figure 1D), $F(4, 36) = 3.576$, $p = 0.015$, corresponding to the amount of learning in the training data in these conditions. The magnitude of AULCSF improvement in the high-reward condition was significantly greater than those in the block- ($p = 0.013$), low- ($p = 0.014$), and no-reward ($p = 0.004$) conditions in post hoc least significant difference (LSD) tests. The magnitudes of improvement in the low- and no-reward conditions were not significantly different ($p = 0.808$), nor did block reward differed from no reward ($p = 0.682$). Subliminal reward led to a larger improvement of AULCSF than no, low, or block reward ($p = 0.031$, $p = 0.076$, $p = 0.080$) and did not differ significantly from that in the high-reward condition ($p = 0.564$). AULCSF improvements in the trained eye were significantly correlated with the total payment (Pearson's $R = 0.746$, $p < 0.001$; Supplementary Figure S1); so were the AULCSF improvements in the untrained eye ($R = 0.531$, $p < 0.001$).

The observers in the five reward conditions did not differ before training as seen in a $5 \times 2 \times 7$ ANOVA performed on the pretraining CSFs with reward condition (no, low, block, subliminal, and high) as a between-subjects factor and eye (trained and untrained)

and spatial frequency as within-subject factors. The main effect of reward condition was not significant, $F(4, 36) = 1.496$, $p = 0.224$ (Supplementary Figure S2A), indicating that the pretraining CSFs were comparable across the five groups. In addition, a one-way ANOVA was performed on the cutoff frequencies used in the training phase. There was no significant difference among the groups (no, low, block, subliminal, and high: 24.42 ± 1.16, 23.43 ± 1.29, 23.67 ± 1.18, 22.17 ± 1.25, 22.06 ± 1.14 c/°; $F(4, 36) = 0.721$, $p = 0.583$.

Generalization of perceptual learning

To quantify the generalizability to other untrained spatial frequencies, we computed the bandwidth of perceptual learning, that is, the bandwidth of the difference between the posttraining and pretraining CSFs (Huang et al., 2008). The bandwidth of perceptual learning in the no-, low-, block-, subliminal-, and high-reward conditions was 2.73 ± 0.62, 2.59 ± 0.23, 2.75 ± 0.52, 3.28 ± 0.22, and 4.40 ± 0.41 octaves, respectively, $F(4, 36) = 3.881$, $p = 0.010$. The bandwidth of perceptual learning in the high-reward condition was significantly broader than those in the no- ($p = 0.015$), low- ($p = 0.019$), and block- ($p = 0.010$) reward conditions but comparable to that in the subliminal-reward condition ($p = 0.172$) in post hoc LSD tests.

Perceptual learning also transferred to the untrained eye with 0.25 ± 0.11, 0.46 ± 0.15, 0.69 ± 0.31, 0.75 ± 0.16, and 1.33 ± 0.29 $\log_{10}$ units of AULCSF improvements in the no-, low-, block-, subliminal-, and high-reward conditions in the untrained eye. A one-way ANOVA with reward condition as a between-subjects factor revealed that the improvements of AULCSF in the untrained eye differed significantly among groups, $F(4, 36) = 3.605$, $p = 0.014$. LSD post hoc tests revealed that the magnitude of improvements in the high-reward condition was (marginally) significantly greater than those in the no- ($p = 0.001$), low- ($p = 0.012$), block- ($p = 0.045$), and subliminal- ($p = 0.099$) reward conditions although there was no significant difference among the last four conditions (all ps > 0.1). In the high-reward condition, the magnitude of the AULCSF improvements in the trained eye was significantly greater than that in the untrained eye (1.92 ± 0.28 vs. 1.33 ± 0.29), $t(7) = 2.612$, $p = 0.035$.

Defining the eye-transfer index as the ratio between the magnitudes of AULCSF improvements in the untrained and trained eyes, we found that the eye-transfer index was comparable across reward conditions (0.45 ± 0.24, 0.51 ± 0.21, 0.73 ± 0.42, 0.49 ± 0.18, and 0.71 ± 0.18; all ps > 0.1).

Retention

Retention of perceptual learning (Zhou et al., 2006) was assessed on seven of the eight observers in the high-reward group at least 5 months after training. The observers retained their AULCSF improvements by 68.2% and 77.4% in the trained and untrained eyes, respectively (see Figure 1E and Supplementary Figure S3).

Discussion

The type and magnitude of reward in the five compound reward manipulations (no, block, low, subliminal, and high reward) were effective in substantially modulating the rate of perceptual learning of contrast detection. There was some learning in all the reward conditions although the rate of learning was much faster in the high-reward group (power function rates of 0.09 ± 0.01 and 0.74 ± 0.01 in the no- and high-reward groups, respectively). The high-reward group, which received trial-by-trial reward signals as well as the block and session signals seen by the block-reward group, clearly generated the fastest learning, achieving the highest magnitude of improvement by the end of training. These reward-induced differences in the rate of learning occurred even though response accuracy feedback, and so information, was equated in all reward conditions.

Generalization of training was assessed by comparing pretraining and posttraining assessments of the CSF, summarized by the AULCSF in the trained and untrained eyes. The perceptual learning induced by training at the cutoff frequency to some degree generalized both across spatial frequencies with the bandwidth of the pretraining to posttraining change of the AULCSF in the trained eye being higher in the high-reward condition. And, although observers were unaware of its existence, the subliminal trial-by-trial reward led to a faster learning and a larger improvement of AULCSF in the trained eye than block reward even though the two conditions appeared the same to observers. As assessed by the eye-transfer index, transfer to the untrained eye occurred (proportionally) for all the reward conditions.

Several pairings of reward conditions offer controlled comparisons. For example, the difference between the high-reward condition, which yields rapid learning, and the low-reward condition, which yields much slower learning, is the combination of base pay and conversion rate for performance-based rewards: learning is much faster when compensation depends on performing well and slower when compensation is largely determined by base pay. Learning is significantly faster when the trial-by-trial reward messages are easily visible than when they are subliminal although trial-by-trial subliminal

reward is surprisingly effective. On the other hand, block reward, in which observers see only both block- and session-level reward messages, is far less effective when compared with high trial-by-trial reward and, at the same time, exceeding that of no reward.

Reward has been hypothesized to operate on perceptual learning via direct modulation or through elevation of attention and/or arousal (Della Libera & Chelazzi, 2006; Peck, Jangraw, Suzuki, Efem, & Gottlieb, 2009). Although this design cannot definitively disentangle the direct effects of rewards from possible indirect influences of reward via attention or arousal mechanisms, some of the comparisons strongly suggest that the role of rewards likely involves more than modulations of attention or arousal. For example, trial-by-trial and block conditions yield substantially the same reward and total compensation because both include block and session rewards and use the same conversion rule, yet the high-reward condition is more effective. To the degree that attention and arousal are related to overall compensation, the degree of arousal should be about comparable in these two conditions. These results suggest that the enhanced perceptual learning induced by trial-by-trial high monetary reward is unlikely to be simply explained by elevated arousal or attention.

In summary, Experiment 1 demonstrated that monetary reward, and especially high monetary reward, can enhance visual perceptual learning and transfer in contrast detection—even when trial-by-trial feedback on response accuracy equates the information provided to the observer.

Experiment 2: High reward boosts perceptual learning through enhanced internal noise reduction

Experiment 2 used the external-noise method and observer-model approach (Dosher & Lu, 1998, 1999; Lu & Dosher, 2008) to analyze the mechanisms of perceptual learning in high- and no-reward conditions in the contrast-detection task. The perceptual template model (PTM; Figure 2A) was used to assess changes of observer inefficiency, including internal noise, template gain, and system nonlinearity (Dosher & Lu, 1998, 1999; Lu & Dosher, 2008). The PTM uses manipulations of external noise in the stimulus to determine how the observer changes due to perceptual learning (Dosher & Lu, 1998). This experiment trains in zero external noise in either the high-reward or no-reward conditions of Experiment 1 and evaluates changes between pretesting and posttesting in the CSF measured in either zero or high external noise in the nondominant eye.

Methods

Observers

Eleven observers (23.55 ± 0.67 years) with normal or corrected-to-normal vision and informed written consent participated in this experiment. None were aware of the purpose of the study. The work was carried out in accordance with the Declaration of Helsinki.

Design

Observers participated in pretraining and posttraining assessments and training in either the *high-* ($n = 6$) or the *no-* ($n = 5$) reward conditions. All visual tasks were performed with the nondominant eye. Unless specified, the procedures were the same as in Experiment 1. Before training, the monocular CSF was measured in the zero and high external-noise conditions (one session with two blocks of 100 trials in each external-noise condition) using the qCSF method (Hou et al., 2010; Lesmes, Lu, Baek, & Albright, 2010) in which test stimuli of different spatial frequencies had a fixed number of cycles (see below).

In the training task, observers were trained in a sine-wave grating detection task (eight sessions, each with seven blocks of 80 trials) at their individual cutoff spatial frequency, estimated from the pretraining CSF in the zero external-noise condition only. Monocular CSFs in the two external-noise conditions were reassessed after training. Vertical sine-wave gratings at 10 spatial frequencies (0.5, 0.67, 1, 1.3, 2, 2.67, 4, 5.3, 8, and 16 c/°) and three full cycles were used to measure the CSF. The size of the gratings was inversely proportional to their spatial frequencies and subtended 6°, 4.5°, 3°, 2.25°, 1.5°, 1.125°, 0.75°, 0.5625°, 0.375°, and 0.1875° at a viewing distance of 138 cm. External noise images were constructed from Gaussian distributed pixel contrasts with $\mu = 0$ and $\sigma = 0$ and 0.24 in the zero and high external-noise conditions, respectively. The size of the external-noise elements was scaled with that of the signal grating to maintain 15 noise elements per grating to maintain a constant signal and external-noise spectral relationship across different spatial frequency conditions. The signal and external-noise images were combined through temporal integration with a sequence of five 35.3-ms frames: two external-noise frames, one signal or blank frame, and two additional external-noise frames. All external-noise frames consisted of independently sampled contrasts.

Data analysis

Motivated by principles in signal processing and neurophysiology, Lu and Dosher (1999, 2008) developed the PTM that characterizes human performance
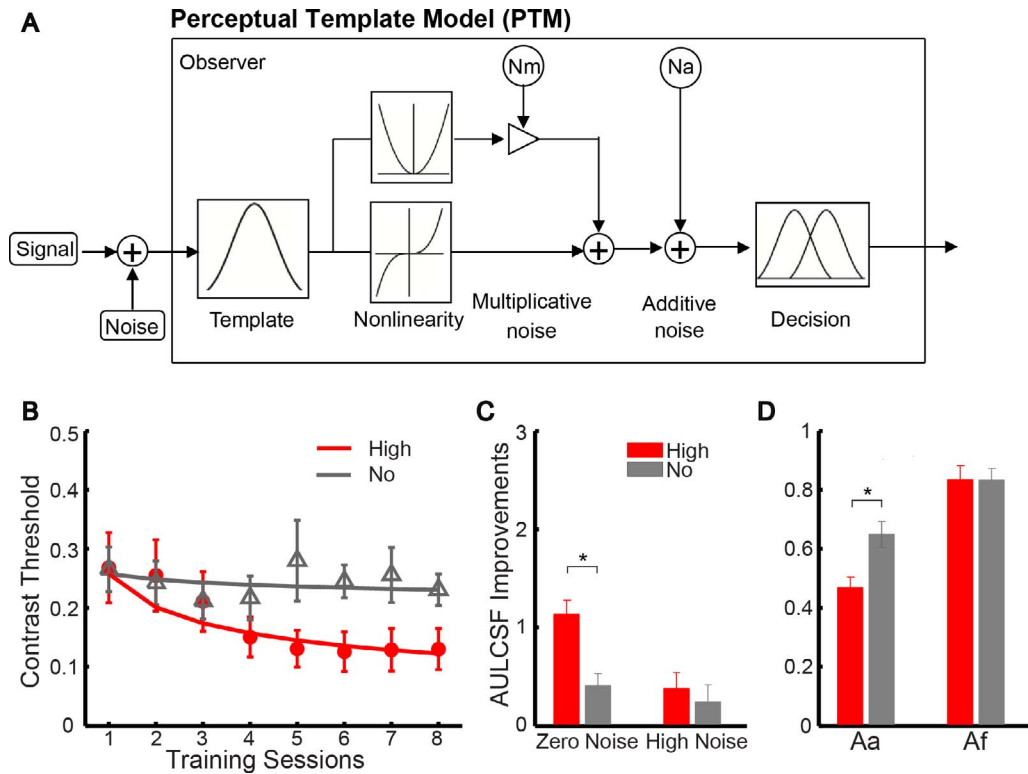
Figure 2. (A) The PTM. The PTM contains five main components: (a) a perceptual template, (b) nonlinear transducer, (c) a multiplicative internal noise source, (d) an additive internal noise source, and (e) a decision process. (B) Learning curves—contrast threshold as a function of training session—were fit by a power function, $C(t) = C_0 t^{-\rho}$, where C_0 is the initial threshold, t is the training session, and ρ is the learning rate. The high- and no-reward learning curves differed only in their learning rates. Red and gray symbols and lines represent data and model fits in the high- and no-reward conditions, respectively. (C) Improved AULCSF in the zero and high external noise conditions. Red and gray bars denote high- and no-reward conditions. (D) A_a and A_f in high (red bar) and no (gray bar) reward conditions fitted by the perceptual template model (1 = no improvement). Error bars indicate standard error.

in perceptual tasks in terms of perceptual template(s), transducer nonlinearity, internal additive noise, and internal multiplicative noise (Figure 2A). The performance of an observer, d', is expressed as

$$d' = \frac{(\beta c)^{\gamma}}{\sqrt{N_{ext}^{2\gamma} + N_{mul}^2\left((\beta c)^{2\gamma} + N_{ext}^{2\gamma}\right) + N_{add}^2}}, \quad (3)$$

where c is signal contrast, β is the gain on a signal valued stimulus processed through the template, γ characterizes the system's nonlinearity, N_{add} is the standard deviation of the internal additive noise, N_{ext} is the standard deviation of external noise, and N_{mul} is the proportional constant of multiplicative noise.

In the PTM framework, perceptual learning improves performance in one or more of three different mechanisms: (a) stimulus enhancement that amplifies both the signal and external noise. This is mathematically equivalent to reducing internal additive noise by a factor of A_a ($0 < A_a \leq 1$), (b) template retuning that reduces external noise by a factor of A_f ($0 < A_f \leq 1$), and (c) multiplicative noise reduction that reduces the

constant for multiplicative noise by a factor of A_m ($0 < A_m \leq 1$). Equation 3 can be rewritten to incorporate the learning parameters as follows:

$$d' = \frac{(\beta c)^{\gamma}}{\sqrt{\left(A_f N_{ext}\right)^{2\gamma} + A_m^2 N_{mul}^2\left((\beta c)^{2\gamma} + \left(A_f N_{ext}\right)^{2\gamma}\right) + \left(A_a N_{add}\right)^2}}. \quad (4)$$

In the PTM framework, slope constancy of the psychometric functions in all external noise conditions before and after training indicates that the multiplicative noise and nonlinearity remained constant before and after training (Lu & Dosher, 1999, 2008). Given this slope constancy before and after training (see Slope check in Supplementary Information), we only considered a model with learning parameters A_a and A_f.

Results

Sensitivity to reward and punishment

The SRPQ revealed no significant difference in reward seeking between the two groups (11.50 ± 1.18

vs. 13.67 ± 1.29), $t(9) = -1.204$, $p = 0.259$, but slightly higher punishment avoidance for the no-reward group (8.50 ± 1.43 vs. 14.60 ± 1.17), $t(9) = -3.209$, $p = 0.011$. In this case, lower punishment avoidance in the high-reward group than that in the no-reward group would weaken, not amplify, any difference found between the two groups.

Pretraining CSF comparison

The observers in the high- and no-reward groups had comparable performance before training in the CSF measures, which were submitted to a $2 \times 2 \times 10$ analysis of variance with reward condition as a between-subjects factor and external-noise level (0 and 0.24) and spatial frequency (0.5, 0.67, 1, 1.3, 2, 2.67, 4, 5.3, 8, and 16 c/°) as within-subject factors (Supplementary Figure S2B). No significant main effect of reward condition was found, $F(1, 9) = 0.932$, $p = 0.360$, indicating that the initial performance of the two groups was comparable. In addition, the cutoff frequencies used in the training phase showed no significant difference between the two groups (10.30 ± 0.52 vs. 10.65 ± 0.54 c/°), $t(9) = -0.460$, $p = 0.656$).

Learning curves

The learning curves for the high- and no-reward conditions, shown in Figure 2B, were analyzed using a procedure similar to that in Experiment 1. Consistent with Experiment 1, training of contrast detection in zero external noise with high reward led to a faster learning rate than that with no reward (0.34 ± 0.03 vs. 0.06 ± 0.02). Training also led to greater improvements of AULCSF in the zero external-noise condition (1.13 ± 0.14 vs. 0.40 ± 0.12), $t(9) = 3.764$, $p = 0.004$, but not in the high external-noise condition (0.37 ± 0.15 vs. 0.24 ± 0.17), $t(9) = 0.579$, $p = 0.577$.

Mechanisms of the enhanced learning

PTM analysis of the CSFs in both the zero and high external-noise conditions (Chen et al., 2014; Yan et al., 2015) identified a mixture of an internal noise-reduction mechanism (to $46.89\% \pm 3.49\%$ and $64.87\% \pm 4.37\%$ of the pretraining levels in the high- and no-reward conditions, respectively, averaged across spatial frequencies; A_a in Figure 2D) and an external noise-exclusion mechanism (to $83.45\% \pm 4.78\%$ and $83.35\% \pm 3.86\%$ of the pretraining levels in the high- and no-reward conditions, respectively, averaged across spatial frequencies; A_f in Figure 2D; see PTM fitting in Supplementary Information for details). High reward enhanced internal noise reduction, $t(9) = -3.253$, $p = 0.010$, but the magnitudes of external-noise exclusion

were comparable in the two reward groups, $t(9) = 0.015$, $p = 0.988$.

Discussion

This experiment examined the mechanisms of visual perceptual learning in the contrast-detection task using an external noise manipulation and the framework of the PTM (Dosher & Lu, 1998, 1999; Lu & Dosher, 1999, 2008). The results suggest that training with high reward boosted the rate and magnitude of perceptual learning through enhanced internal noise reduction relative to training with no reward, and training with either high or no reward led to smaller and approximately equivalent improvements in external-noise exclusion.

Experiment 3: Effects of compound monetary reward on Vernier offset judgments

Experiment 3 examined whether the titration of learning by reward extended to training a quite different task: Vernier offset hyperacuity (Fahle & Edelman, 1993; Herzog & Fahle, 1997; McKee & Westheimer, 1978; Saarinen & Levi, 1995; Xiao et al., 2008) or judging the horizontal offsets (in arcs of visual angle) between two vertically stacked Gabor sine-wave patches. Vernier offset judgment tasks, like contrast-detection tasks, are thought to involve representations in early visual areas. Learning in Vernier offset judgments, almost always using feedback without explicit rewards, is generally specific to the spatial/retinal location of training (Fahle, 1997; Fahle & Morgan, 1996).

In this experiment, Vernier offsets were measured in upper left (Loc1) and lower right (Loc2) peripheral locations in pretests and posttests. Training was performed in the upper left quadrant. Observers identified letters briefly presented at fovea to guarantee fixation. Generalization was assessed in the untrained compared to the trained location. The high- and no-reward conditions were identical to those of Experiments 1 and 2 except that criteria for absolute performance rewards were modified for Vernier performance.

Methods

Observers

Sixteen observers (23.37 ± 0.74 years) with normal or corrected-to-normal vision and informed written
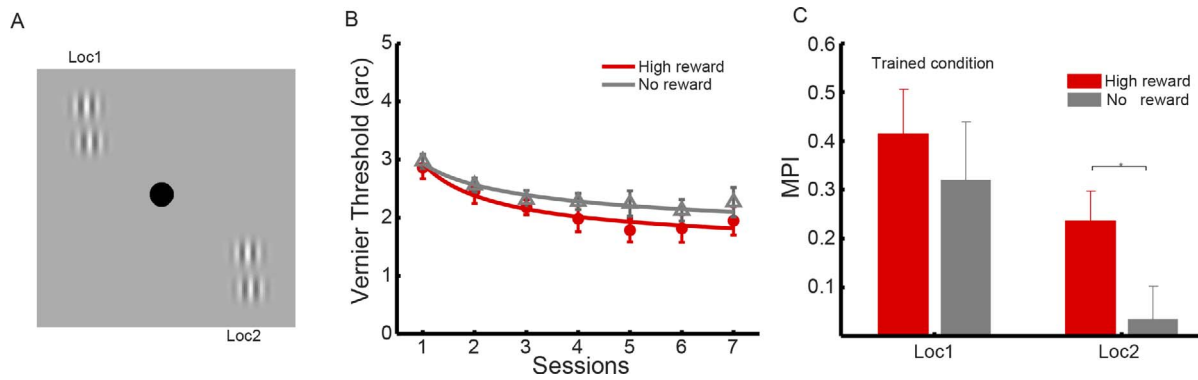
Figure 3. (A) Stimulus configuration in a Vernier discrimination task (Experiment 3). The stimulus was presented in either the upper left (Loc1) or lower right (Loc2) visual quadrant. Training was carried out at the upper left quadrant with vertical orientation (Loc1). (B) Learning curves in the high-reward (red) and no-reward (gray) conditions. Vernier thresholds over sessions were fitted with power functions. (C) Mean Percentage Improvement (MPI) of Vernier thresholds in trained and untrained locations.

consent participated in this study. None were aware of the purpose of the study. The work was carried out in accordance with the Declaration of Helsinki.

Design

Observers participated in pretraining and posttraining assessments and training in the high-reward ($n = 8$) and no-reward ($n = 8$) conditions. Before training, threshold Vernier offsets were measured in two locations (Loc1: upper left vs. Loc2: lower right) in one session (two blocks of 100 trials) before training. Then observers were trained in Vernier offset judgments in Loc1 for five sessions (each with seven blocks of 80 trials). Threshold offsets in two conditions (Loc1 vs. Loc2) were reassessed after training.

Procedure

Observers were asked to make two judgments after each trial: first to report the foveal letter (H or N) for fixation control and then to report the offset direction of the Vernier stimulus. Vernier stimulus was presented at either the upper left (Loc1) or lower right (Loc2) visual quadrant (Figure 3A). Observers judged whether the lower Gabor was to the left or right of the upper Gabor. Only the performance in the Vernier task was related to reward points in the high-reward condition. As in Experiment 1, observers in the high-reward conditions gained reward points when their performance improved between trials, blocks, and sessions during the training phase (see Experiment 1, Methods). The only change was that the additional 10,000, 20,000, and 30,000 points were awarded if the observer's Vernier threshold was reduced by 40%, 50%, and 60% compared to the pretest values. No reward or feedback is provided during pretesting and

posttesting. Additionally, observers filled out the SRPQ.

Procedure

A fixation point (0.2°) preceded each trial by 400 ms. Then, the Vernier stimulus (contrast = 0.45, SF = 3 c/°, and σ = 0.29°) was presented at 5° retinal eccentricity (positional jitter = 0.25°) for 200 ms, during which time a sequence consisted of nine small black letters appeared at fixation, and observers identified a target letter. The purpose of the foveal letter report was to guarantee fixation. Observers viewed the display monocularly at a viewing distance of 1.38 m. The offset threshold was assessed by a three-down/one-up staircase with a step size of 10% (of the current offset) that converges to 79.4% correct. Auditory feedback was given on correct responses during training, not during pretest or posttest assessment of threshold.

Data analysis

Improvements in Vernier threshold across two test and five training sessions were fit with an elaborated power function:

$$C(t) = C_0 t^{-\rho} + \alpha, \quad (5)$$

where $C(t)$ is the Vernier threshold in the tth session, C_0 is the initial Vernier threshold, ρ is the learning rate, and α is the intercept or lower asymptote. For the two groups of observers receiving no or high reward, the complete model has six parameters ($2C_0$, 2ρ, 2α). The model that postulates no effects of reward ($1C_0$, 1ρ, 1α) assumes that C_0, ρ, and α are the same for the two groups. Between the fully saturated model and the no-change model, a lattice of models with different numbers of parameters was explored.

The mean percentage improvement (MPI), which compares performance in the posttraining assessment to the pretraining baseline, was calculated by the following function:

$$MPI = \frac{(threshold\ in\ pretest - threshold\ in\ posttest)}{threshold\ in\ pretest}. \quad (6)$$

This measure can be applied to both the trained location and the untrained location, where it is a measure of transfer.

Results

Sensitivity to reward and punishment

There was no significant difference between the two groups of observers in the high- and no-reward conditions in reward sensitivity, $t(14) = 0.703$, $p = 0.494$, or punishment sensitivity, $t(14) = 0.514$, $p = 0.615$.

Learning

Figure 3B shows the learning curves, threshold offset for the Vernier judgment as a function of training (in seconds of arc), for the two reward conditions. Similar to Experiments 1 and 2, the reduced model that assumed different learning rates but the same initial threshold and intercept ($1C_0$, 2ρ, 1α) fit the data best. With four parameters, the fit of this reduced model was statistically equivalent to the full model ($r^2 = 94.29\%$ vs. 95.58%), $F(2, 8) = 1.168$, $p = 0.358$, and significantly superior to the most reduced model (94.29% vs. 40.10%), $F(1, 10) = 25.341$, $p < 0.001$. The parameters of the best model are $C_0 = 1.22$, $\alpha = 1.88$, and $\rho = 0.44$ (± 0.06), and 2.73 (± 0.87) in the no- and high-reward conditions, respectively. The presence of high reward significantly increased the learning rate (Figure 3B).

The high-reward condition yielded somewhat, but not significantly, higher MPI in the trained location (comparing pretraining to posttraining; MPI, see Equation 6; high vs. no: 41% ± 9% vs. 32% ± 12%), $t(14) = 0.631$, $p = 0.538$ (Figure 3C).

Generalization

The MPI of the high-reward group was significantly larger than that of the no-reward group in the untrained location (Loc2; high vs. no: 24% ± 6% vs. 3% ± 7%), $t(14) = 2.200$, $p = 0.045$. At the same time, the transfer index for retinal location (the MPI at the untrained location divided by the MPI at the trained location) was not significantly different in the high- versus no-reward conditions (high vs. no: 0.48 ± 0.15 vs. 0.27 ± 0.05, $p = 0.191$), reflecting equivalent

transfer as a proportion of learning. These results indicated that high monetary reward speeded up the learning rate of Vernier offset judgment and improved transfer to an untrained location without changing the interlocation transfer index.

Central letter identification

The performance in the central letter identification task was comparable between the two groups before (high vs. no: 96.9% vs. 95.3%), $t(14) = 1.290$, $p = 0.218$, and after training (high vs. no: 97.1% vs. 97.7%), $t(14) = -0.549$, $p = 0.592$. Letter identification received only auditory feedback during the course of training. The task achieved a high level of accuracy even in pretesting and was unaffected by training on the Vernier task.

Discussion

This experiment examined the effectiveness and generalizability of learning in a Vernier offset task with composite monetary rewards compared with no reward in the presence of informational feedback. As in the previous experiments, high monetary reward speeded up the rate of learning. It also improved transfer to an untrained location without changing the interlocation transfer index. The MPI in the untrained location was near zero, replicating the frequent finding of little location transfer in tasks trained with feedback only; in contrast, the MPI in the untrained location was significant when training used composite reward.

Experiment 4: Effect of monetary reward on global motion direction discrimination

The contrast-detection and Vernier offset judgment tasks are thought to involve representations in early visual areas (Duncan & Boynton, 2003; Ress, Backus, & Heeger, 2000). In this experiment, we investigated whether high-compound monetary reward also influences learning and transfer in a midlevel visual task, e.g., global motion-direction discrimination (Koyama et al., 2005; Vaina, Belliveau, Des Roziers, & Zeffiro, 1998). In previous studies, perceptual learning improved motion-direction discrimination along a trained cardinal direction, and the learning failed to transfer to an untrained direction (Ball & Sekuler, 1982). The previous studies either used feedback in the absence of explicit reward or small rewards equivalent to the low-reward condition in the current paper.
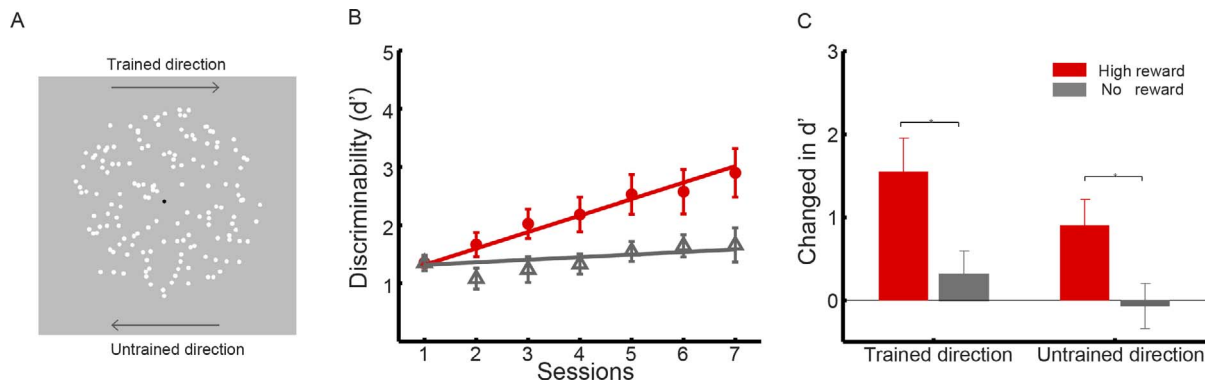
Figure 4. (A) Stimulus configuration in a global motion direction–discrimination task (Experiment 4). The circular fixation remained stationary while the dots moved in a single direction (0° or 180°). A two-interval, forced-choice paradigm was used. Observers were asked to judge whether the motion directions of the two stimuli were the same or different. (B) Learning curves in the high-reward (red) and no-reward (gray) conditions. Discrimination sensitivity (d') over sessions was fitted with linear functions. (C) Improvements of d' in trained (0°) and untrained (180°) directions. Error bars indicate standard error.

Methods

Observers

Fifteen observers (23.38 ± 0.76 years) with normal or corrected-to-normal vision and informed written consent participated in this study. None were aware of the purpose of the study. The work was carried out in accordance with the Declaration of Helsinki.

Design

Observers participated in pretraining and posttraining assessments and training in the high-reward ($n = 7$) and no-reward ($n = 8$) conditions. The task was to judge whether global motion directions in two intervals were the same or different (by a small angle). The dependent measure was d' in performing this same–different judgment. Accuracy of the judgments was assessed for two reference motion directions (0° and 180°) in one session (four blocks of 64 trials) before and after training. During training, observers were trained in one reference direction (0°) for five sessions; each contained seven blocks of 80 trials. Similar to Experiment 1, observers in the high-reward conditions gained reward points when their performance improved between trials, blocks, and sessions during the training phase. Here, the criteria for additional reward points for absolute performance were set at 80%, 85%, 90%, and 95% correct. Additionally, observers completed the SRPQ.

Procedure

Four hundred dots (0.18° × 0.18°) moved along a single direction within a circular aperture of 8° in diameter with a speed of 10°/s. In each trial, the reference (0° or 180°) and test (0° ± 2.5° or 180° ± 2.5°) were separately presented in two 500-ms stimulus intervals in a random order separated by a 200-ms interstimulus interval. A small dark fixation point (0.15°) was always present in the center of the display (Figure 4A). Observers viewed the display binocularly with a distance of 0.6 m. The observers judged whether dots in the two intervals moved in the same direction or not. Auditory feedback was given on correct responses during training in both reward conditions. The delivery of rewards was identical to that of Experiment 1 except for adjusting the absolute reward criteria as described above.

Data analysis

Discrimination scores in the global motion direction–discrimination task across two test and five training sessions were fit with a linear function (Dosher & Lu, 2005):

$$d'(t) = \rho t + \alpha, \quad (7)$$

where $d'(t)$ is the discriminability score in the tth session, α is the intercept, and ρ is the learning rate. The d' was calculated as the difference of z score between hit and false alarm rates. The improvements from pretraining to posttraining assessments in the global motion task were assessed by the difference in d' performance or $\Delta d'$.

Results

Sensitivity to reward and punishment

Sensitivity to reward, $t(13) = -1.610$, $p = 0.131$, and punishment avoidance, $t(13) = -0.224$, $p = 0.826$, were comparable in the two groups.

Learning

Figure 4B shows the learning curves with high reward and no reward in the global motion task, measured as improvements in discriminability *d'*. As in the previous experiments, a model with different learning rates for the two reward conditions provided a significantly better fit than a model assuming the same learning rate, $r^2 = 94.82\%$ vs. 23.82%, $F(1, 11) = 150.916$, $p < 0.001$. The parameters of the best-fitting model with different learning rates for the high reward and no reward conditions are $\alpha = 1.06$, $\rho = 0.06$ (± 0.01) and 0.31 (± 0.01) in the no- and high-reward conditions, respectively. High composite reward increased the linear learning rate in global motion-direction discrimination relative to no reward.

Generalization

Figure 4C shows the increase in *d'* in posttraining relative to pretraining assessments of global motion-direction discrimination. The magnitude of learning in the high-reward condition exceeded that of the no-reward condition in the trained direction, high vs. no: 1.54 ± 0.41 vs. 0.31 ± 0.29 improvement in *d'*, $t(13) = 2.512$, $p = 0.026$. This result also extended to improvements in the untrained global motion direction, high vs. no: 0.90 ± 0.31 vs. −0.07 ± 0.27 improvements in *d'*, $t(13) = 2.345$, $p = 0.036$. Again, the transfer index (*d'* improvement in the untrained direction divided by the *d'* improvement in the trained direction) between the two reward conditions was comparable (high vs. no: 0.69 ± 0.17 vs. 0.49 ± 0.38, $p = 0.632$).

Discussion

Global motion direction discrimination, thought to be a midlevel visual task (Duncan & Boynton, 2003; Ress et al., 2000), showed the same improvements in both learning and generalization with high composite monetary reward compared to no reward. The no-reward group shows little generalization to the untrained global motion direction, similar to results in other studies in the literature using feedback but no explicit reward. Training with high composite monetary rewards demonstrates significant generalization in the untrained direction, and the proportional transfer index of untrained to trained directions is equivalent.

General discussion

In a series of experiments, we systematically investigated the effects and mechanisms of compound monetary reward on the magnitude, rate, and transfer of perceptual learning using a reward structure that consisted of trial-by-trial, between-block, and between-session rewards. All reward conditions included feedback on response accuracy, separating the role of information in learning from the effects of reward. The effects of no, low, block, subliminal, and high reward on perceptual learning and transfer were fully assessed in a monocular contrast-detection task. The compound reward manipulations created reward incentive structures that ranged from base pay that is independent of performance rewards in the no-reward condition to a situation in which the large majority of monetary rewards depend upon performance. Together, this study has, for the first time, provided a form of dose–response assessment of the effects of reward on the behavioral improvements in perceptual learning.

The rates of learning for trained spatial frequency in the five reward conditions of Experiment 1—no, low, block, subliminal, and high reward—were 0.09 ± 0.01, 0.15 ± 0.01, 0.31 ± 0.01, 0.45 ± 0.02, and 0.74 ± 0.01, respectively. The highest learning rate occurred in the high trial-by-trial reward condition, second was the subliminal trial-by-trial reward condition, and then the block reward condition. Perceptual learning in the high trial-by-trial reward condition also improved contrast sensitivity over a broad range of spatial frequencies for both the trained and untrained eyes. Effectively, the more that is learned in the primary trained task, the more learning there is to transfer, leading to a comparable transfer ratio. Furthermore, the effects of learning were long lasting. In Experiment 2, an external noise manipulation and analysis by the PTM showed that training with high reward compared with no reward improved the rate and magnitude of perceptual learning primarily through enhanced internal noise reduction although both high- and no-reward groups showed a common slight improvement in external noise exclusion. Experiments 3 and 4 in Vernier offset and global motion discrimination showed that the effects of high reward also extended to a range of perceptual learning tasks.

Together, these results suggest that high monetary reward, especially with trial-by-trial rewards as well as block and session rewards, can play a significant role in enhancing the magnitude and transfer of perceptual learning without changing interocular, interlocation, and interdirection transfer ratios. These results may have important implications for our understanding of the nature of visual plasticity as well as practical applications of perceptual learning (Goldstone, 1998).

It is not surprising that no and low reward—given equivalent response accuracy feedback—led to the least amount of relative performance improvements (Huang et al., 2008). In the no-reward condition of Experiment 1, significant learning was found in six out of 11 subjects, and the average contrast sensitivity improve-

ment at the trained frequency was 6.48 dB. In an earlier study that employed similar experimental setups and targeted the difference in learning characteristics in normal and amblyopic groups (Huang et al., 2008), we found significant learning effects of contrast detection in the no-reward condition in nine out of 14 subjects and an average improvement of 5.8 dB at the trained frequency, comparable to the results in the current study.

Reward, attention, and arousal have highly overlapping brain circuitries (Maunsell, 2004; Schultz, 2006). In these studies, high reward might have improved perceptual performance and perceptual learning either through direct reward circuits or through associated changes in arousal and/or top-down attention (Della Libera & Chelazzi, 2006; Peck et al., 2009). Indeed, some other researchers have also suggested that perceptual learning may improve performance through improved attention and/or arousal (Ahissar & Hochstein, 1997; Gilbert, Sigman, & Crist, 2001; Xiao et al., 2008). The high-reward condition consisted of trial-by-trial rewards that may provide a trial-by-trial differential signal between predicted and received reward on the scale of seconds (Nomoto, Schultz, Watanabe, & Sakagami, 2010; Schultz, 1998), and between-block and between-session rewards may generate elevated arousal or top-down attention on the scale of minutes or hours (Roesch & Olson, 2007). If perceptual learning in the high-reward condition was driven only by improved top-down attention or arousal, one might expect the same amount of performance improvement in the trained and untrained eyes because top-down attention and arousal operate binocularly (Karni & Sagi, 1991; Schwartz, Maquet, & Frith, 2002). This was not the case. In fact, we found that the magnitude of the improvements in the AULCSF in the trained eye was significantly greater than that in the untrained eye in the high-reward condition in Experiment 1. The result suggests that perceptual learning in the high-reward condition was not driven solely by improved top-down attention or arousal. In the block-reward condition, no trial-by-trial reward was available, but observers improved their performance more than in the no- and low-reward conditions. The partial but lower efficacy of block reward parallels the results with block compared to trial-by-trial feedback (Herzog & Fahle, 1997; Liu et al., 2014; Shibata, Yamagishi, Ishii, & Kawato, 2009). We speculate that block reward may influence learning through self-assessed aspects of performance (Liu et al., 2014) or through changes in arousal or effort. Trial-by-trial subliminal reward enhanced perceptual learning only slightly less than the high-reward condition, consistent with other findings that unconscious reward could improve perceptual learning (Xue et al., 2015). Finally, explicit trial-by-trial high reward generated the greatest amount of perceptual learning, suggesting that conscious awareness of reward may amplify its effects slightly and further enhance perceptual learning (Zedelius, Veling, & Aarts, 2012). Taken together, these results suggest the structure of reward itself is the largest determinant of perceptual learning and transfer.

In Experiments 3 and 4, we examined the generalizability of the effects of high reward structure in Vernier offset judgment and global motion-direction discrimination, two tasks that usually exhibit high degrees of location and/or direction specificity (Ball & Sekuler, 1982; Poggio, Fahle, & Edelman, 1992). Interestingly, high monetary reward facilitated the learning rate and improved location transfer in Vernier offset judgment and also increased the rate and magnitude of learning and direction transfer in global motion-direction discrimination. In addition, the transfer index, which measures how much of what is learned is transferred, was comparable among high- and no-reward groups both in Vernier offset judgment and global motion discrimination. Because contrast detection, Vernier offset judgment, and global motion discrimination may involve a wide range of visual cortical areas (Furmanski & Engel, 2000; Ress et al., 2000; Vaina et al., 1998), our results suggest that high monetary reward could impact neural plasticity in different stages of visual processing, and the effects of high reward may extend to a wide range of perceptual learning tasks.

The relative transfer to the untrained eye, motion direction, and retinal location was largely determined by how much was learned in the training task. The compound reward structure greatly enhanced the rate and magnitude of learning in the trained condition and improved absolute performance in these tasks in the untrained eye, motion direction, and retinal locations. This reward paradigm may be practical and attractive.

The finding that training with high reward improved perceptual learning differentially through enhanced internal noise reduction sheds light on the mechanism through which reward enhances perceptual learning. The reward circuits interact with brain regions associated with cognition and motor control (Haber & Knutson, 2010). Traditionally, reward has been shown to affect the late stages of visual processing, including visual-motor transformation (Schultz, Tremblay, & Hollerman, 2000), decision making (Hampton & O'Doherty, 2007), and overt behavior (Behrens, Woolrich, Walton, & Rushworth, 2007). It has also been shown that stimuli associated with high reward often induced better performance, larger event-related potential amplitude, and stronger fMRI signals in the reward system (Krawczyk, Gazzaley, & D'Esposito, 2007; Pessiglione et al., 2007). More recently, a number of studies have also showed that reward could affect

early sensory processing (Seitz et al., 2009; Serences, 2008).

In this study, we found significant generalization of learning to other untrained spatial frequencies and the untrained eye (Experiment 1), to untrained retinal locations (Experiment 3), and to untrained global motion directions (Experiment 4) in the high-reward conditions even while the no-reward conditions often showed more specificity. Some theories of perceptual learning have argued that the degree of generalization negatively relates to the hierarchy level of learning (Ahissar & Hochstein, 2004; Shibata, Sagi, & Watanabe, 2014). In the integrated reweighting theory (Dosher et al., 2013), learning at the level of invariant representations leads to generalization. One possibility, then, is that reward-induced noise reduction may occur at a relatively late stage, at least after binocular combination and at the level of largely spatial frequency–invariant representations (Duncan & Boynton, 2003; H.-H. Li, Rankin, Rinzel, Carrasco, & Heeger, 2017; Ress et al., 2000).

One idea is that reward processing involves the dopamine pathways and a convergence of several corticostriatal projections (Arias-Carrion & Poppel, 2007). Animal studies on primates found that the rise in the dopamine concentration of the basal ganglia reaches its peak about 1 s after the onset of the reward-related stimulus, starts to decline after 2 s, and drops down to the baseline level after about 4 s. The trial-by-trial high monetary reward in our study may induce the firing of dopamine neurons to promote reward-seeking behaviors and, therefore, change the characteristics of learning (Ariansen et al., 2012; Schluter, Mitz, Cheer, & Averbeck, 2014; Yoshimi et al., 2011). Additionally or alternatively, reward may act within the span of a single trial to improve the signal-to-noise ratio of sensory representations in task-relevant channels in early sensory areas through direct projections from the basal forebrain to/on primary visual cortex or through indirect modulations of activities in early visual cortical areas (Baldassi & Simoncini, 2011; Bhattacharyya, Veit, Kretz, Bondar, & Rainer, 2013). A recent study found that the primary reward (water) given during training reactivated the reward system and its interaction with perceptual processing during subsequent REM sleep (Berard, Barnes-Diana, Nanez, Sasaki, & Watanabe, 2015). The reactivation of the reward circuitry during REM sleep may strengthen and consolidate visual perceptual learning (Karni, Tanne, Rubenstein, Askenasy, & Sagi, 1994; Sasaki et al., 2010).

From a modeling point of view, reward could enhance perceptual learning in two possible ways. One direct way that reward may affect perceptual learning is through reward prediction error, i.e., the difference between the expected and actual reward in a given trial, which is a major component in reinforcement learning algorithms (Dayan & Balleine, 2002). The other, indirect way that reward could enhance perceptual learning is through improved sensory encoding or decision making. For example, in augmented Hebbian learning, the rate of learning is determined by the product of sensory signal and decision (Petrov et al., 2005). Improving either component could improve perceptual learning. Both reinforcement and Hebbian rules as well as some forms of hybrid rules have been proposed for perceptual learning (Law & Gold, 2009; Petrov et al., 2005). In this study, we found that, although reward has a primary role in learning, attention and arousal could also contribute to enhanced perceptual learning. Our results are consistent with a hybrid learning rule with both reinforcement and Hebbian components in which prediction error and improved sensory and decision processes can all contribute to enhanced perceptual learning (Petrov et al., 2005).

We used a type of secondary reward, i.e., monetary compensation, in the current study. It should be noted that other forms of reward, including endogenous rewards associated with the primary task (Xue et al., 2015), primary rewards, such as water (Seitz et al., 2009) and juice (Imai, Kim, Sasaki, & Watanabe, 2014), and social reward (Hayward, Pereira, Otto, & Ristic, 2018), have also been found to be effective in initiating and/or improving perceptual learning. Whether there is a single mechanism for the different forms of reward or there are multiple reinforcement processes that differentially modulate visual perceptual learning remains to be elucidated.

In summary, we found that trial-by-trial high monetary reward boosted the rate, magnitude, and generalizability of perceptual learning; that high monetary reward differentially enhanced internal noise-reduction mechanisms of perceptual learning; and the effect was universal in a range of tasks. This suggests that high reward may be an important component in applications of perceptual learning. Theoretically, feedback, arousal/attention, and reward could all contribute to the enhancement of perceptual learning in a hybrid-learning rule that incorporates effects of reward, attention, and feedback and both Hebbian and reinforcement learning components.

Keywords: perceptual learning, reward, transfer, perceptual template model

Acknowledgments

Commercial relationships: Z-LL owns intellectual property rights on adaptive testing technologies and has equity interest in Adaptive Sensory Technology, Inc. FH, Z-LL, and C-BH have equity interest in Jiangsu Juehua Medical Technology Co., LTD.
Corresponding author: Chang-Bing Huang; Zhong-Lin Lu; Barbara Anne Dosher.
Email: huangcb@psych.ac.cn; lu.535@osu.edu; bdosher@uci.edu.
Address: Chinese Academy of Sciences, Beijing, China; Laboratory of Brain Processes (LOBES), Center for Cognitive and Brain Sciences, Center for Cognitive and Behavioral Brain Imaging, and Departments of Psychology, The Ohio State University, Columbus, OH, USA; Department of Cognitive Sciences and Institute of Mathematical Behavioral Sciences, University of California, Irvine, CA, USA.

References

Ahissar, M., & Hochstein, S. (1997, May 22). Task difficulty and the specificity of perceptual learning. *Nature*, *387*(6631), 401–406.

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464.

Ariansen, J., Heien, M. L. A. V., Hermans, A., Phillips, P. E. M., Hernadi, I., Bermudez, M., . . . Wightman, R. M. (2012). Monitoring extracellular pH, oxygen, and dopamine during reward delivery in the striatum of primates. *Frontiers in Behavioral Neuroscience*, *6*(36), 1–10.

Arias-Carrion, O., & Poppel, E. (2007). Dopamine, learning, and reward-seeking behavior. *Acta Neurobiologiae Experimentalis*, *67*(4), 481–488.

Baldassi, S., & Simoncini, C. (2011). Reward sharpens orientation coding independently of attention. *Frontiers in Neuroscience*, *5*, 13.

Ball, K., & Sekuler, R. (1982, November 12). A specific and enduring improvement in visual motion discrimination. *Science*, *218*(4573), 697–698.

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221.

Berard, A., Barnes-Diana, T., Nanez, J., Sasaki, Y., & Watanabe, T. (2015). External reward facilitates visual perceptual learning over a night's sleep [Abstract]. *Journal of Vision*, *15*(12): 1302, https://doi.org/10.1167/15.12.1302. [Abstract]

Bhattacharyya, A., Veit, J., Kretz, R., Bondar, I., & Rainer, G. (2013). Basal forebrain activation controls contrast sensitivity in primary visual cortex. *BMC Neuroscience*, *14*(1), 55.

Brainard, D. (1997). Psychophysics software for use with MATLAB. *Spatial Vision*, *10*, 433–436.

Chen, G., Hou, F., Yan, F.-F., Zhang, P., Xi, J., Zhou, Y., . . . Huang, C.-B. (2014). Noise provides new insights on contrast sensitivity function. *PloS One*, *9*(3), e90579.

Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, *36*(2), 285–298.

Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science*, *17*(3), 222–227.

Dosher, B., & Lu, Z.-L. (2017). Visual perceptual learning and models. *Annual Review of Vision Science*, *3*(1), 343–363.

Dosher, B. A., Jeter, P., Liu, J. J., & Lu, Z. L. (2013). An integrated reweighting theory of perceptual learning. *Proceedings of the National Academy of Sciences, USA*, *110*(33), 13678–13683.

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, USA*, *95*(23), 13988–13993.

Dosher, B. A., & Lu, Z.-L. (1999). Mechanisms of perceptual learning. *Vision research*, *39*(19), 3197–3221.

Dosher, B. A., & Lu, Z.-L. (2005). Perceptual learning in clear displays optimizes perceptual expertise: Learning the limiting process. *Proceedings of the National Academy of Sciences, USA*, *102*(14), 5286–5290.

Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38*(4), 659–671.

Fahle, M. (1997). Specificity of learning curvature, orientation, and Vernier discriminations. *Vision Research*, 37(14), 1885–1895.

Fahle, M., & Edelman, S. (1993). Long-term learning in Vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research*, 33(3), 397–412.

Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, 6(3), 292–297.

Fahle, M., & Poggio, T. (2002). *Perceptual learning*. Cambridge, MA: MIT Press.

Furmanski, C. S., & Engel, S. A. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience*, 3(6), 535–536.

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, 31(5), 681–697.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49(1), 585–612.

Haber, S. N., & Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology*, 35(1), 4–26.

Hampton, A. N., & O'Doherty, J. P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proceedings of the National Academy of Sciences, USA*, 104(4), 1377–1382.

Hayward, D. A., Pereira, E. J., Otto, A. R., & Ristic, J. (2018). Smile! Social reward drives attention. *J Exp Psychol Hum Percept Perform*, 44(2), 206–214.

Herzog, M. H., & Fahle, M. (1997). The role of feedback in learning a Vernier discrimination task. *Vision Research*, 37(15), 2133–2141.

Hou, F., Huang, C.-B., Lesmes, L., Feng, L.-X., Tao, L., Zhou, Y.-F., & Lu, Z.-L. (2010). qCSF in clinical application: Efficient characterization and classification of contrast sensitivity functions in amblyopia. *Investigative Ophthalmology & Visual Science*, 51(10), 5365–5377.

Huang, C.-B., Zhou, Y., & Lu, Z.-L. (2008). Broad bandwidth of perceptual learning in the visual system of adults with anisometropic amblyopia. *Proceedings of the National Academy of Sciences, USA*, 105(10), 4068–4073.

Imai, H., Kim, D., Sasaki, Y., & Watanabe, T. (2014). Reward eliminates retrieval-induced forgetting. *Proceedings of the National Academy of Sciences*, 111(48), 17326–17329.

Jeter, P. E., Dosher, B. A., Liu, S.-H., & Lu, Z.-L. (2010). Specificity of perceptual learning increases

with increased training. *Vision Research*, 50(19), 1928–1940.

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J. J., & Sagi, D. (1994, July 29). Dependence on REM sleep of overnight improvement of a perceptual skill. *Science*, 265(5172), 679–682.

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, USA*, 88(11), 4966–4970.

Kennerley, S. W., & Wallis, J. D. (2009). Reward-dependent modulation of working memory in lateral prefrontal cortex. *The Journal of Neuroscience*, 29(10), 3259–3270.

Kim, Y.-H., Kang, D.-W., Kim, D., Kim, H.-J., Sasaki, Y., & Watanabe, T. (2015). Real-time strategy video game experience and visual perceptual learning. *The Journal of Neuroscience*, 35(29), 10485–10492.

Koop, M. R., Applegate, R. A., & Howland, H. C. (1996). Changes in the area under the log contrast sensitivity function (AULCSF) with myopic refractive error. *Investigative Ophthalmology & Visual Science*, 37(3), 1482–1482.

Koyama, S., Sasaki, Y., Andersen, G. J., Tootell, R. B., Matsuura, M., & Watanabe, T. (2005). Separate processing of different global-motion structures in visual cortex is revealed by FMRI. *Current Biology*, 15(22), 2027–2032.

Krawczyk, D. C., Gazzaley, A., & D'Esposito, M. (2007). Reward modulation of prefrontal and visual association cortex during an incentive working memory task. *Brain Research*, 1141, 168–177.

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.

Lesmes, L. A., Lu, Z.-L., Baek, J., & Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: The quick CSF method. *Journal of Vision*, 10(3):17, 1–21, https://doi.org/10.1167/10.3.17. [PubMed] [Article]

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49(2), 467–477.

Li, H. H., Rankin, J., Rinzel, J., Carrasco, M., & Heeger, D. J. (2017). Attention model of binocular rivalry. *Proceedings of the National Academy of Sciences of the United States of America*, 114(30), E6192–E6201.

Li, X., Lu, Z.-L., Xu, P., Jin, J., & Zhou, Y. (2003).

Generating high gray-level resolution monochrome displays with conventional computer graphics cards and color monitors. *Journal of Neuroscience Methods*, *130*(1), 9–18.

Liu, J., Dosher, B., & Lu, Z.-L. (2014). Modeling trial by trial and block feedback in perceptual learning. *Vision Research*, *99*, 46–56.

Liu, J., Dosher, B. A., & Lu, Z.-L. (2015). Augmented Hebbian reweighting accounts for accuracy and induced bias in perceptual learning with reverse feedback. *Journal of Vision*, *15*(10):10, 1–21, https://doi.org/10.1167/15.10.10. [PubMed] [Article]

Liu, J., Lu, Z.-L., & Dosher, B. A. (2010). Augmented Hebbian reweighting: Interactions between feedback and training accuracy in perceptual learning. *Journal of Vision*, *10*(10):29, 1–14, https://doi.org/10.1167/10.10.29. [PubMed] [Article]

Liu, J., Lu, Z. L., & Dosher, B. A. (2012). Mixed training at high and low accuracy levels leads to perceptual learning without feedback. *Vision Research*, *61*(61), 15–24.

Lu, Z.-L., & Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America*, *16*(3), 764–778.

Lu, Z.-L., & Dosher, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychological Review*, *115*(1), 44–82.

Lu, Z.-L., & Dosher, B. A. (2009). Mechanisms of perceptual learning. *Learning & Perception*, *1*(1), 19–36, https://doi.org/10.1556/LP.1.2009.1.3.

Lu, Z.-L., Hua, T., Huang, C.-B., Zhou, Y., & Dosher, B. A. (2011). Visual perceptual learning. *Neurobiology of Learning and Memory*, *95*(2), 145–151.

Maunsell, J. H. (2004). Neuronal representations of cognitive state: Reward or attention? *Trends in Cognitive Sciences*, *8*(6), 261–265.

McKee, S. P., & Westheimer, G. (1978). Improvement in Vernier acuity with practice. *Perception & Psychophysics*, *24*(3), 258–262.

Nomoto, K., Schultz, W., Watanabe, T., & Sakagami, M. (2010). Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *The Journal of Neuroscience*, *30*(32), 10692–10702.

Pascucci, D., & Turatto, M. (2013). Immediate effect of internal reward on visual adaptation. *Psychological Science*, *24*(7), 1317–1322.

Peck, C. J., Jangraw, D. C., Suzuki, M., Efem, R., & Gottlieb, J. (2009). Reward modulates attention

independently of action value in posterior parietal cortex. *The Journal of Neuroscience*, *29*(36), 11182–11191.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007, May 11). How the brain translates money into force: A neuroimaging study of subliminal motivation. *Science*, *316*(5826), 904–906.

Petrov, A. A., Dosher, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review*, *112*(4), 715–743.

Petrov, A. A., Dosher, B. A., & Lu, Z. L. (2006). Perceptual learning without feedback in non-stationary contexts: Data and model. *Vision Research*, *46*(19), 3177–3197.

Poggio, T., Fahle, M., & Edelman, S. (1992, May 15). Fast perceptual learning in visual hyperacuity. *Science*, *256*(5059), 1018–1021.

Ress, D., Backus, B. T., & Heeger, D. J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. *Nature Neuroscience*, *3*(9), 940–945.

Roesch, M. R., & Olson, C. R. (2007). Neuronal activity related to anticipated reward in frontal cortex. *Annals of the New York Academy of Sciences*, *1121*(1), 431–446.

Saarinen, J., & Levi, D. M. (1995). Perceptual learning in Vernier acuity: What is learned? *Vision Research*, *35*(4), 519–527.

Sagi, D. (2011). Perceptual learning in vision research. *Vision Research*, *51*(13), 1552–1566.

Sasaki, Y., Nanez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience*, *11*(1), 53–60.

Sasaki, Y., Náñez, J. E., & Watanabe, T. (2012). Recent progress in perceptual learning research. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 293–299.

Schluter, E. W., Mitz, A. R., Cheer, J. F., & Averbeck, B. B. (2014). Real-time dopamine measurement in awake monkeys. *PloS One*, *9*(6), e98692.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*(1), 1–27.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*, 87–115.

Schultz, W., Tremblay, L., & Hollerman, J. R. (2000).

Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex*, *10*(3), 272–283.

Schwartz, S., Maquet, P., & Frith, C. (2002). Neural correlates of perceptual learning: A functional MRI study of visual texture discrimination. *Proceedings of the National Academy of Sciences, USA*, *99*(26), 17137–17142.

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, *61*(5), 700–707.

Seitz, A. R., Yamagishi, N., Werner, B., Goda, N., Kawato, M., & Watanabe, T. (2005). Task-specific disruption of perceptual learning. *Proceedings of the National Academy of Sciences, USA*, *102*(41), 14895–14900.

Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, *60*(6), 1169–1181.

Shibata, K., Sagi, D., & Watanabe, T. (2014). Two-stage model in perceptual learning: Toward a unified theory. *Annals of the New York Academy of Sciences*, *1316*, 18–28.

Shibata, K., Yamagishi, N., Ishii, S., & Kawato, M. (2009). Boosting perceptual learning by fake feedback. *Vision Research*, *49*(21), 2574–2585.

Skinner, B. F. (1938). *The behavior of organisms: an experimental analysis*. Oxford, England: Appleton-Century.

Vaina, L. M., Belliveau, J. W., Des Roziers, E. B., & Zeffiro, T. A. (1998). Neural systems underlying learning and representation of global motion. *Proceedings of the National Academy of Sciences, USA*, *95*(21), 12657–12662.

Watanabe, T., Náñez, J. E., & Sasaki, Y. (2001, October 25). Perceptual learning without perception. *Nature*, *413*(6858), 844–848.

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.

Xue, X., Zhou, X., & Li, S. (2015). Unconscious reward facilitates motion perceptual learning. *Visual Cognition*, *23*(1–2), 161–178.

Yan, F.-F., Zhou, J., Zhao, W., Li, M., Xi, J., Lu, Z.-L., & Huang, C.-B. (2015). Perceptual learning improves neural processing in myopic vision. *Journal of Vision*, *15*(10):12, 1–14, https://doi.org/10.1167/15.10.12. [PubMed] [Article]

Yoshimi, K., Naya, Y., Mitani, N., Kato, T., Inoue, M., Natori, S., . . . Kitazawa, S. (2011). Phasic reward responses in the monkey striatum as detected by voltammetry with diamond microelectrodes. *Neuroscience Research*, *71*(1), 49–62.

Zedelius, C. M., Veling, H., & Aarts, H. (2012). When unconscious rewards boost cognitive task performance inefficiently: The role of consciousness in integrating value and attainability information. *Frontiers in Human Neuroscience*, *6*, 219.

Zhou, Y., Huang, C., Xu, P., Tao, L., Qiu, Z., Li, X., & Lu, Z.-L. (2006). Perceptual learning improves contrast sensitivity and visual acuity in adults with anisometropic amblyopia. *Vision Research*, *46*(5), 739–750.