Two-stage perceptual learning to break visual crowding

Ziyun Zhu*

Department of Psychology and Beijing Key Laboratory of Behavior and Mental Health
Key Laboratory of Machine Perception (Ministry of Education)
Peking-Tsinghua Center for Life Sciences
PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, P.R. China

Zhenzhi Fan*

Department of Psychology and Beijing Key Laboratory of Behavior and Mental Health
Key Laboratory of Machine Perception (Ministry of Education)
Peking-Tsinghua Center for Life Sciences
PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, P.R. China

Fang Fang

Department of Psychology and Beijing Key Laboratory of Behavior and Mental Health
Key Laboratory of Machine Perception (Ministry of Education)
Peking-Tsinghua Center for Life Sciences
PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, P.R. China

When a target is presented with nearby flankers in the peripheral visual field, it becomes harder to identify, which is referred to as crowding. Crowding sets a fundamental limit of object recognition in peripheral vision, preventing us from fully appreciating cluttered visual scenes. We trained adult human subjects on a crowded orientation discrimination task and investigated whether crowding could be completely eliminated by training. We discovered a two-stage learning process with this training task. In the early stage, when the target and flankers were separated beyond a certain distance, subjects acquired a relatively general ability to break crowding, as evidenced by the fact that the breaking of crowding could transfer to another crowded orientation, even a crowded motion stimulus, although the transfer to the opposite visual hemi-field was weak. In the late stage, like many classical perceptual learning effects, subjects' performance gradually improved and showed specificity to the trained orientation. We also found that, when the target and flankers were spaced too finely, training could only reduce, rather than completely eliminate, the crowding effect. This two-stage

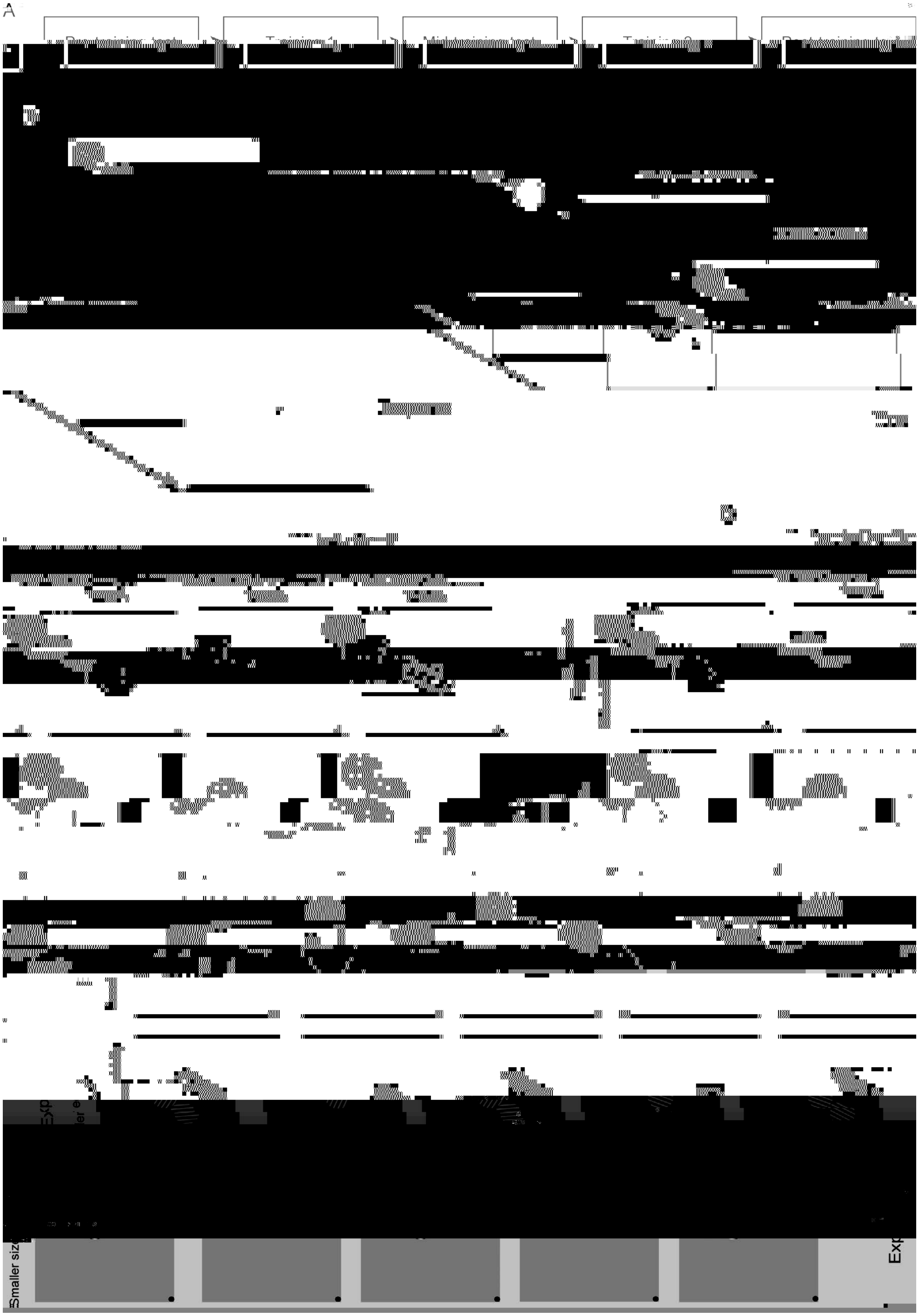

Figure 1. Experimental protocol and stimuli. (A) Each experiment consisted of five phases: pretraining test (Pre), Training 1, mid-training test (Mid), Training 2, and posttraining test (Post). (B) Schematic description of a two-alternative forced-choice trial in a QUEST staircase for measuring the orientation discrimination threshold with a crowded target. (C) Trained and test stimuli in Experiments 1–5. Black dots represent the fixation point. The stimuli were presented in the upper-left visual quadrant, except that the isolated and crowded untrained targets in Experiment 3 were presented in the upper-right visual quadrant.

were asked to make a two-alternative forced-choice judgment of the orientation of the second target relative to the first one (clockwise or counterclockwise). Informative feedback was provided after each response by brightening (correct response) or dimming (wrong response) the fixation point briefly, which facilitated learning (Goldhacker, Rosengarth, Plank, & Greenlee, 2014). The next trial began 800 to 1200 ms after feedback. $\Delta\theta$ varied trial by trial and was controlled by QUEST staircases to estimate subjects' discrimination thresholds at 75% correct.

During the three test phases, subjects' orientation discrimination thresholds were measured with four test stimuli: the crowded trained target, the isolated trained target, the crowded untrained target, and the isolated untrained target (Figure 1C, first row). The untrained target was identical to the trained target except that its orientation was perpendicular to that of the trained target. Thirty-two QUEST staircases (same as above), eight for each test stimulus, were completed in a random order. Starting values in the QUEST staircases were identical. During Training 1, subjects continued practicing with the crowded (trained) target until the mean threshold from five consecutive QUEST staircases was lower than the threshold measured with the isolated trained target at Pre. During Training 2, subjects underwent six more daily training sessions with the crowded target.

Experiments 2 and 3 had the same design and trained stimulus as Experiment 1. Two of the four test stimuli (the crowded trained target and the isolated trained target) in Experiment 1 were also used in Experiments 2 and 3. In Experiment 2, the gratings in the crowded untrained target and the isolated untrained target in Experiment 1 were replaced with random-dot kinematograms (RDKs; radius: 1.5°; dot density: $8/°^2$; velocity: 10°/s; luminance: 0.01 cd/m^2). The moving direction of the target RDK deviated from the orientation of the trained target in Experiment 1 by 60°, either clockwise or counterclockwise. The directions of two flanker RDKs were randomized (Figure 1C, second row). Similar to the orientation discrimination measurement, we measured subjects' motion direction discrimination thresholds with these two new test stimuli. In Experiment 3, the crowded trained target and the isolated trained target in Experiment 1 were also presented in the upper-right visual quadrant, referred to as the crowded untrained target and the isolated untrained target, respectively (Figure 1C, third row).

Experiment 4 also had the same design as Experiment 1. The trained and test stimuli in Experiment 4 were similar to those in Experiment 1, except that the stimuli were presented at 6° eccentricity and the radius of the target and flanker gratings was reduced to 0.98° according to the cortical magnification factor for matching the cortical representation sizes of the stimuli between Experiments 1 and 4 (Duncan & Boynton, 2003). The center-to-center distance between the target and the flankers was 1.96° (Figure 1C, fourth row).

The design of Experiment 5 was slightly different from that of Experiment 1. It had only three phases: Pre, Training 1, and Mid. The contrast and spatial frequency of the target and flankers were identical to those in Experiment 1, but their radius and center-to-center distance were reduced to half of those in Experiment 1. The stimuli were presented at the same eccentricity as that in Experiment 1 (Figure 1C, fifth row). During Training 1, all subjects underwent 10 daily training sessions with the crowded target.

Data analysis

For each test stimulus, discrimination thresholds from eight QUEST staircases were averaged as a measure of subjects' discrimination performance at Pre, Mid, and Post. Subjects' performance improvements with a test stimulus from Pre to Mid and from Mid to Post were calculated as (pretraining threshold – mid-training threshold)/pre-training threshold $\times$ 100%) and (mid-training threshold – posttraining threshold)/mid-training threshold $\times$ 100%), respectively. Because the learning effects after training at 67.5° and 157.5° were very similar, they were pooled together for further analysis.

Results

Experiment 1: Perceptual learning with crowded orientation and its transfer to crowded orthogonal orientation

Subjects were trained to perform an orientation discrimination task at 67.5° or 157.5° with a crowded target. Their orientation discrimination thresholds were measured using QUEST staircase throughout training. A daily training session consisted of 30 QUEST staircases of 40 trials. Before training, we measured subjects' orientation discrimination thresholds at Pre with four test stimuli: the crowded trained target, the isolated trained target, the crowded untrained target, and the isolated untrained target (Figure 2A). The orientation of the untrained target was perpendicular to that of the trained target. Subjects' discrimination thresholds were much higher with the crowded targets than with the isolated targets: crowded trained target versus isolated trained target, $t(7) = 10.12$, $p < 0.001$; crowded untrained target versus isolated untrained target, $t(7) = 10.23$, $p < 0.001$, demonstrating that the

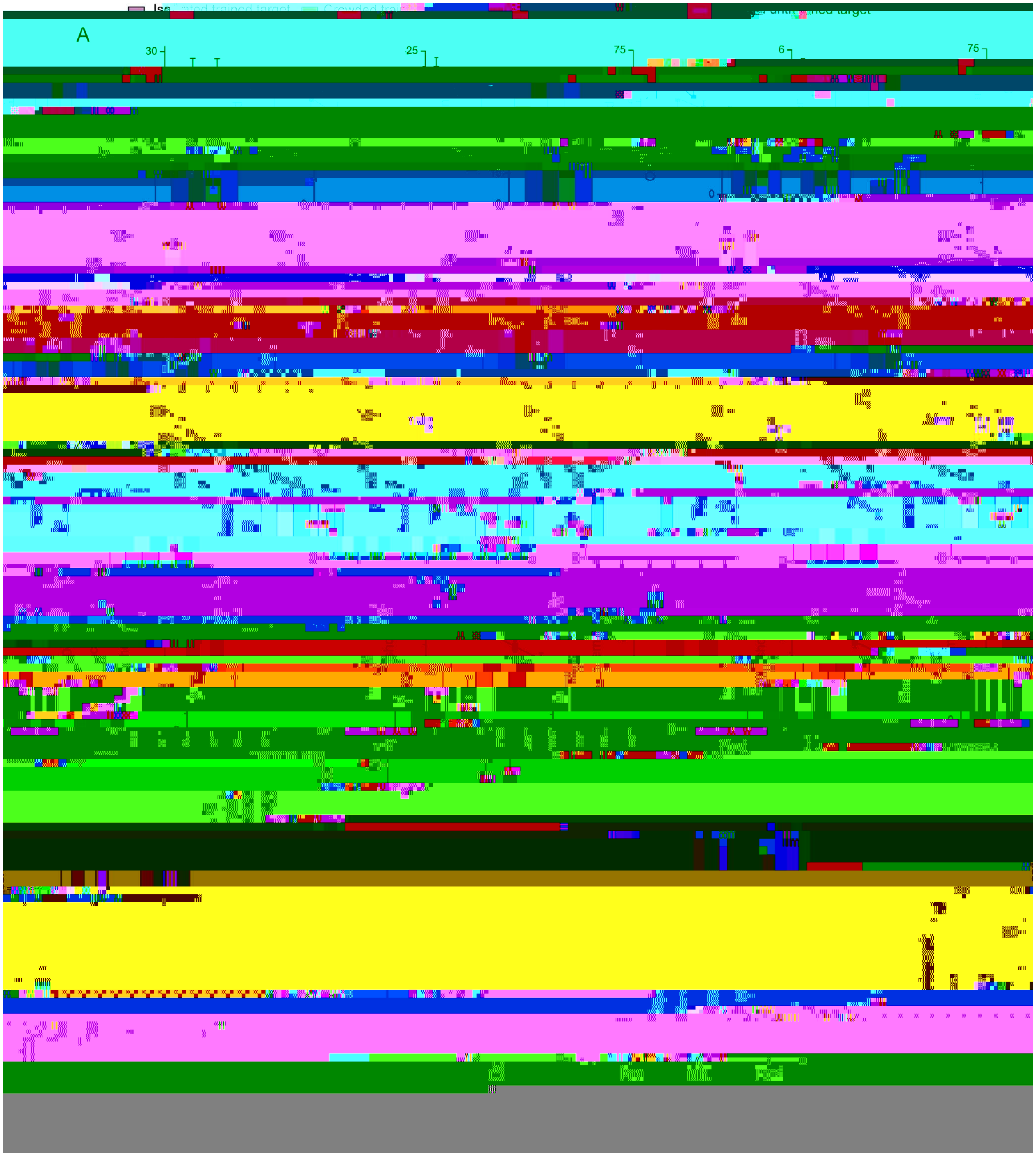

Figure 2. Psychophysical results of Experiments 1–5 and the control experiment. (A–D) First column (from left to right): discrimination thresholds for the four test stimuli at Pre, Mid, and Post. Second column: learning curve during Training 1. For individual subjects, staircases during Training 1 were split into six equally sized bins based on the training progress. The average discrimination threshold in each bin was plotted as a function of bin, referred to as the learning curve. Learning curves were then averaged across subjects. Third column: percentage improvements in discrimination performance from Pre to Mid. Fourth column: learning curve during Training 2. Discrimination thresholds are plotted as a function of training day. Fifth column: percentage improvements in discrimination performance from Mid to Post. (E, F) Discrimination thresholds for the four test stimuli at Pre and Mid. Error bars denote 1 SEM across subjects.

presentation of the nearby flankers led to strong crowding.

During Training 1, subjects' performance improved quickly and substantially. The training ceased after 1,760 ± 302 trials (about 1.5 training sessions, throughout the article, X ± Y indicates the mean ± SEM across subjects), because at that time, the mean

0.05. The difference between the improvements with the isolated trained target and the isolated untrained target was significant, $t(8) = 3.801$, $p < 0.01$. The learning effect with the crowded trained target could largely transfer to the crowded untrained target, despite that the two stimuli consisted of dramatically different components (i.e., oriented grating and RDK). However, the transfers to the isolated trained target and the isolated untrained target were much weaker.

After Training 2, the improvements in discrimination performance from Mid to Post were 49.04% ± 4.11% for the crowded trained target, 49.27% ± 3.67% for the isolated trained target, 7.62% ± 3.02% for the crowded untrained target, and 16.89% ± 5.68% for the isolated untrained target, all $t(8) > 2.55$, $p < 0.05$. The learning effect with the crowded trained target completely transferred to the isolated trained target. But the transfers to the crowded untrained target and the isolated untrained target were weak.

These findings provided further evidence that, in the first learning stage, subjects learned to separate the target and flankers presented at the trained location. The improved segmentation ability persisted despite the fact that the trained and test stimuli (oriented grating vs. RDK) were completely different. In the second learning stage, the learning effect showed specificity to the trained feature, replicating the finding in Experiment 1.

Experiment 3: Perceptual learning with crowded orientation and its transfer to the opposite visual hemi-field

Experiment 3 was designed to examine whether the learned ability to break crowding could generalize to the opposite visual hemi-field. The experiment used the same design and stimuli as Experiment 1, except that the crowded trained target and the isolated trained target in Experiment 1 were also presented in the upper-right visual quadrant, referred to as the crowded untrained target and the isolated untrained target, respectively.

At Pre, the crowding effects were very strong in both visual hemi-fields, both $t(7) > 12.97$, $p < 0.001$ (Figure 2C). Training 1 ceased after subjects practiced 2,090 ± 407 trials. It improved subjects' performance dramatically and removed the crowding effect in the trained (i.e., left) visual hemi-field. Performance improvements from Pre to Mid were 72.77% ± 2.33% for the crowded trained target, 31.32% ± 4.90% for the isolated trained target, 34.54% ± 7.03% for the crowded untrained target, and 21.18% ± 4.08% for the isolated untrained target, all $t(7) > 4.52$, $p < 0.01$. Different from Experiments 1 and 2, the transfer of the learning effect to the crowded untrained target was weak in Experiment 3, which was comparable to the transfer to the isolated trained target and the isolated untrained target. This finding demonstrated that the improved segmentation ability after Training 1 manifested largely at the trained location.

From Mid to Post, the improvements with the crowded trained target (53.48% ± 3.48%), the isolated trained target (44.87% ± 4.66%), and the crowded untrained target (22.78% ± 7.12%) were significant, all $t(7) > 2.92$, $p < 0.05$, but not with the isolated untrained target (12.85% ± 10.16%), $t(7) = 1.52$, $p > 0.05$. Again, this finding demonstrated that the learning effect from Training 2 exhibited specificity for the trained orientation at the trained location.

Experiment 4: Perceptual learning with crowded orientation at smaller eccentricity

Experiment 4 examined whether the results in Experiment 1 could be replicated at 6° eccentricity. The stimuli in Experiment 1 were reduced in size according to the cortical magnification factor and then used in Experiment 4. At Pre, the crowding effects were very strong, both $t(7) > 7.11$, $p < 0.001$ (Figure 2D). Training 1 ceased after subjects practiced 1,720 ± 418 trials. From Pre to Mid, the improvements in discrimination performance were 63.39% ± 2.56% for the crowded trained target, 19.01% ± 5.76% for the isolated trained target, 55.43% ± 3.28% for the crowded untrained target, and 12.55% ± 3.10% for the isolated untrained target, all $t(7) > 3.04$, $p < 0.05$. From Mid to Post, the improvements were 57.20% ± 1.95% for the crowded trained target, 49.14% ± 3.94% for the isolated trained target, 18.00% ± 3.71% for the crowded untrained target, and 28.02% ± 2.72% for the isolated untrained target, all $t(7) > 4.22$, $p < 0.01$. The two-stage learning effects were very similar to those in Experiment 1.

Experiment 5: Limited effect of perceptual learning with crowded orientation

In Experiment 5, the stimulus sizes were reduced to half of those in Experiment 1. The stimuli were still presented at the same eccentricity as that in Experiment 1. We examined whether crowding could be completely eliminated with smaller stimuli. At Pre, crowding effects were too strong to measure subjects' orientation discrimination thresholds with the crowded target (not

During Training 1, subjects learned to perform the discrimination task with the crowded target. However, even after 10 days' training, the crowding effect could not be completely eliminated, and training ceased. At Mid, the orientation discrimination thresholds with the isolated target were not significantly different from those at Pre, both

interaction depending on the distance between the target and flankers may play a major role in determining the magnitude of crowding. Taken together, crowding is determined by the combination of constraints at multiple levels of cortical processing, including low-level cortical interaction and high-level attention.

The performance improvement in the early training stage was largely due to the improved general ability of segmenting the target and flankers, which manifested with the crowding configuration (i.e., the radial configuration) used in the study. Distinct from the early training stage, the improvement in the late training stage was mainly attributed to the perceptual learning effect specific to the trained orientation. The visual system might have learned to refine the neural representation of the trained orientation in sensory areas and/or improve relevant decision-making processes in higher cortical areas (N. Chen et al., 2015; Law & Gold, 2008; Schoups et al., 2001). It is noteworthy that, in the early training stage, there was a

letters: does it improve reading speed? Vision Research, 47, 3150–3159.

Chen, J., He, Y., Zhu, Z., Zhou, T., Peng, Y., Zhang, X., & Fang, F. (2014). Attention-dependent early cortical suppression contributes to crowding. Journal of Neuroscience, 34, 10465–10474.

Chen, N., Bi, T., Zhou, T., Li, S., Liu, Z., & Fang, F. (2015). Sharpened cortical tuning and enhanced cortico-cortical communication contribute to the long-term neural mechanisms of visual motion perceptual learning. Neuroimage, 115, 17–29.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nature Review Neuroscience, 3, 201–215.

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

Ghose, G. M., Yang, T., & Maunsell, J. H. (2002). Physiological correlates of perceptual learning in monkey V1 and V2.

Strasburger, H., Harvey, L. O., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. Perception & Psychophysics, 49, 495–508.

Störmer, V. S., Alvarez, G. A., & Cavanagh, P. (2014). Within-hemifiled competition in early visual areas limits the ability to track multiple objects with attention. Journal of Neuroscience, 34, 11526–11533.

Sun, G. J., Chung, S. T., & Tjan, B. S. (2010). Ideal observer analysis of crowding and the reduction of crowding through learning. Journal of Vision, 10(5): 16, 1–14, doi:10.1167/10.5.16. [PubMed] [Article]

Watanabe, T., & Sasaki, Y. (2015). Perceptual learning: Toward a comprehensive theory. Annual Review of Psychology, 1, 197–221.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. Perception & Psychophysics, 33, 113–120.

Westheimer, G., Shimamura, K., & McKee, S. P. (1976). Interference with line-orientation sensitivity. Journal of the Optical Society of America, 66, 332–338.

Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. Trends in Cognitive Sciences, 15, 160–168.

Xiong, Y., Yu, C., & Zhang, J. (2015). Perceptual learning eases crowding by reducing recognition errors but not position errors. Journal of Vision, 15(11):16, 1-13, doi:10.1167/15.11.16. [PubMed] [Article]