

Perceptual learning eases crowding by reducing recognition errors but not position errors

Ying-Zi Xiong

Department of Psychology and Peking–Tsinghua Center for Life Sciences, Peking University, Beijing, China



Cong Yu

Department of Psychology and Peking–Tsinghua Center for Life Sciences, Peking University, Beijing, China



Jun-Yun Zhang

Department of Psychology and Peking–Tsinghua Center for Life Sciences, Peking University, Beijing, China



When an observer reports a letter flanked by additional letters in the visual periphery, the response errors (the crowding effect) may result from failure to recognize the target letter (recognition errors), from mislocating a correctly recognized target letter at a flanker location (target misplacement errors), or from reporting a flanker as the target letter (flanker substitution errors). Crowding can be reduced through perceptual learning. However, it is not known how perceptual learning operates to reduce crowding. In this study we trained observers with a partial-report task (Experiment 1), in which they reported the central target letter of a three-letter string presented in the visual periphery, or a whole-report task (Experiment 2), in which they reported all three letters in order. We then assessed the impact of training on recognition of both unflanked and flanked targets, with particular attention to how perceptual learning affected the types of errors. Our results show that training improved target recognition but not single-letter recognition, indicating that training indeed affected crowding. However, training did not reduce target misplacement errors or flanker substitution errors. This dissociation between target recognition and flanker substitution errors supports the view that flanker substitution may be more likely a by-product (due to response bias), rather than a cause, of crowding. Moreover, the dissociation is not consistent with hypothesized mechanisms of crowding that would predict reduced positional errors.

Introduction

A letter, when flanked by additional letters, becomes difficult to recognize in peripheral vision (Bouma, 1970;

Flom, Heath, & Takahashi, 1963; Levi, 2008). This crowding effect is considered to be a bottleneck in peripheral vision (Levi, 2008). Crowding is attributed to abnormal integration of features at a stage beyond feature detection (Greenwood, Bex, & Dakin, 2009; Levi, Hariharan, & Klein, 2002; Pelli, Palomares, & Majaj, 2004), or limited attention resolution (S. He, Cavanagh, & Intriligator, 1996; Intriligator & Cavanagh, 2001).

Recent studies show that crowding can be alleviated to some degree by perceptual learning in typical and clinical populations (Chung, 2007; Chung, Li, & Levi, 2012; Y. He, Legge, & Yu, 2013; Huckauf & Nazir, 2007; Hussain, Webb, Astle, & McGraw, 2012; Sun, Chung, & Tjan, 2010). For example, Chung (2007) reported that training improves the averaged letter recognition performance at the trained letter separation, and that the learning transfers to other letter separations. However, as detailed later, several types of errors are involved in crowding. It is unclear how these errors are affected by training.

A flanked target is not only more difficult to recognize (recognition errors, Figure 1), it may also be correctly recognized but mislocated to a flanker position (target misplacement errors, Figure 1; Zhang, Zhang, Liu, & Yu, 2012). Zhang et al. (2012) asked observers to report the identity of the central target of a briefly presented trigram, as is commonly done in crowding studies (partial report), and also to report the entire trigram in order (whole report). In the whole-report paradigm, the target identification rate is independent of the position (order) in which the target is reported. Zhang et al. (2012) and others (Y. He et al., 2013) found that under crowded conditions, the target identity is sometimes correctly reported but in the

Citation: Xiong, Y.-Z., Yu, C., & Zhang, J.-Y. (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.

Types of crowding errors

	Stimulus	I	C	O
Recognition error	Response	I	?	O
Target misplacement error	Response	C	X	O
Flanker substitution error	Response	X	I	O

Figure 1. An illustration of three types of crowding errors. The first row shows a three-letter string consisting of an inner flanker (I), a central target (C), and an outer flanker (O). The inner and outer flankers were left and right flankers in the current study, respectively. The second row shows a recognition error when the observer fails to report C at any of the three stimulus positions in whole report. The third row shows a target misplacement error when the observer reports C at a flanker position. X can be I or a no-show letter in this case. The bottom row shows a flanker substitution error when an observer wrongly reports a flanker letter (I in this example) as C. X can be either C or a no-show letter. Notice that more than one type of error could occur in a single trial.

wrong position. These target misplacement errors are not recognition errors. Rather, target misplacement errors and actual recognition errors together contribute to the report errors in a typical crowding task that uses the partial report as a performance measurement.

Another type of position error in crowding is flanker substitution (Figure 1). When failing to recognize the central letter, the observer will sometimes report a flanking letter (flanker substitution errors; Huckauf & Heller, 2002; Krumhansl, 1977; Strasburger, 2005). Flanker substitution errors are regarded as a possible source of crowding (Huckauf & Heller, 2002; Strasburger, 2005). However, an alternative explanation is that when failing to recognize the central letter, the observer may be biased to report a more visible flanker (Strasburger, 2005). Flanker substitution errors may or may not occur when a report error occurs. On the other hand, flanker substitution errors and target misplacement errors are not mutually exclusive, because both could occur in the same trial (for a detailed analysis, see Zhang et al., 2012).

The current study investigates how training to reduce crowding affects each of the three types of errors, in order to provide a more detailed understanding of the mechanisms underlying the reduction of crowding due to perceptual learning. Moreover, the training data enable us to test different hypotheses regarding flanker substitution errors and crowding per se. If flanker

substitution is a cause of crowding, training would be expected to reduce flanker substitution errors. On the other hand, if flanker substitution occurs only because observers are biased to report a more visible flanker when failing to recognize the target, it is more likely a by-product of crowding and should be unchanged by training.

Our results show that training mainly reduces recognition errors but not target misplacement errors, suggesting that the observers learn to recognize the central target but fail to reduce the position errors with recognized target letters. Our results also show that (normalized) flanker substitution errors are unchanged with reduced crowding. This dissociation suggests that target substitution by a more visible flanker is likely a by-product, rather than a cause, of crowding.

Methods

Observers and apparatus

Twenty observers (undergraduate students in their 20s) with normal or corrected-to-normal vision participated in this study. They were new to psychophysical experiments and were unaware of the purpose of the study. Informed consent was obtained from each observer before data collection. This research adhered to the Declaration of Helsinki.

The stimuli were generated with the MATLAB toolbox Psychtoolbox-3 (Pelli, 1997) and presented on a 21-in. Sony G520 color monitor (1600 × 1200 resolution, 0.25 mm × 0.25 mm pixel size, 85-Hz frame rate, 0.66 cd/m² minimum and 92.22 cd/m² maximum luminance). An Eyelink II eye tracker (SR Research, Kanata, Ontario, Canada) was used to monitor eye movements. Viewing was monocular with one eye covered with a translucent plastic pad. (We often compare performance among amblyopic eyes, non-amblyopic fellow eyes, and typical eyes in various studies. So we typically use monocular viewing even if no observers with amblyopia are involved, just in case we need normal data in the future.) A chin-and-head rest was used to keep the head fixed. The viewing distance was 0.8 m. Experiments were run in a dimly lit room.

Stimuli

The test stimuli were meaningless horizontal trigrams (strings of three Sloan letters) presented at 10° retinal eccentricity on the horizontal meridian of the right visual field (Figure 2a). A fixation cross was constantly centered on the monitor screen during the

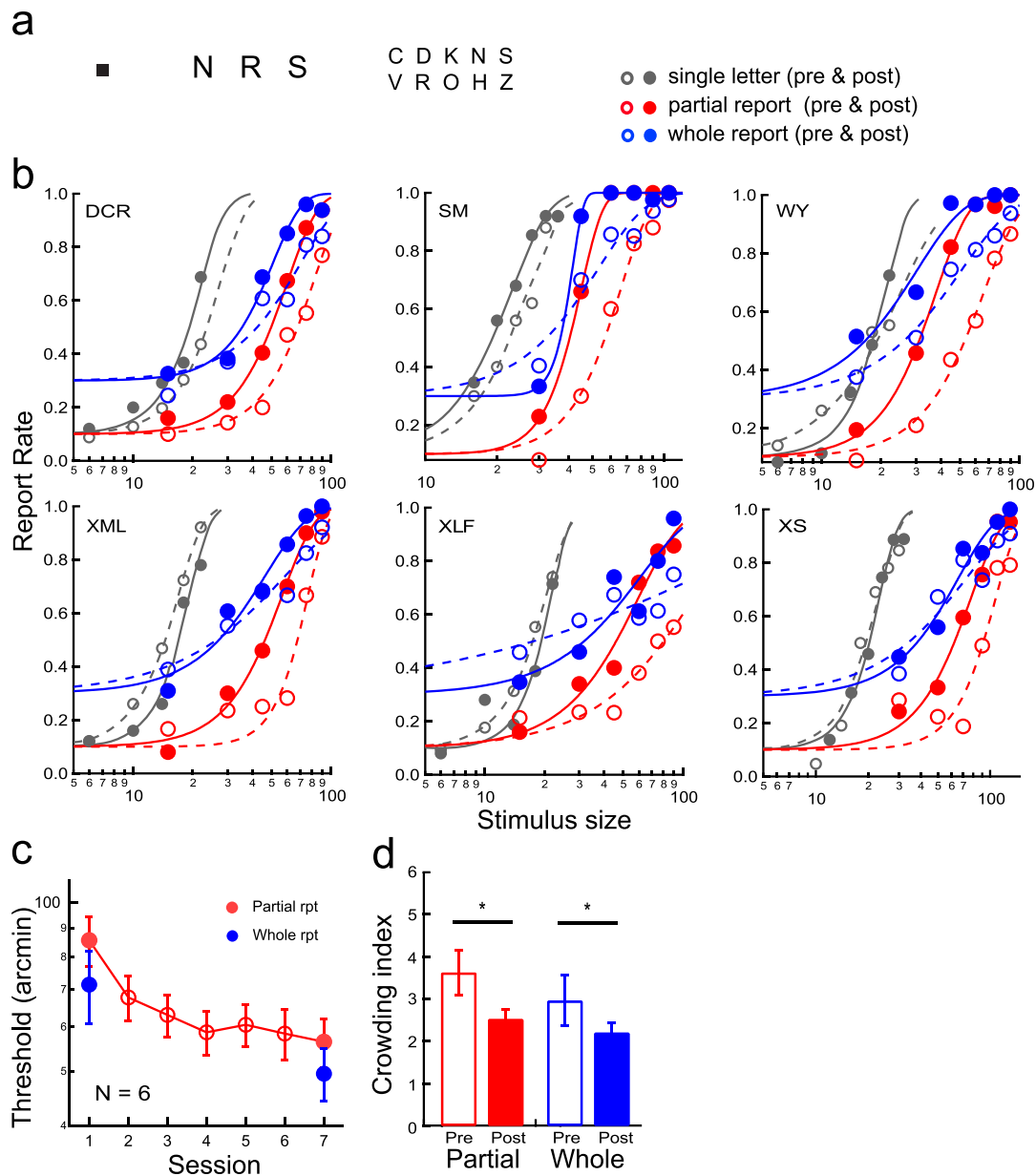


Figure 2. The effects of partial-report training on crowding. (a) Stimuli. Left: a horizontal three-letter string presented in the right visual field. Right: the 10 Sloan letters. (b) Individual psychometric functions for single-letter recognition, partial report, and whole report before and after training. Empty and solid symbols represent pre- and posttraining data, respectively. Dashed and solid curves are Weibull fittings. (c) Letter recognition thresholds estimated with partial-report data through the course of partial-report training, as well as pre- and posttraining thresholds estimated with whole-report data. Here and later, error bars indicate one standard error of the mean. (d) Pre- and posttraining crowding indices estimated with partial- and whole-report data.

experimental session. The three Sloan letters, which were black on a full-luminance white screen, were randomly selected from a list of 10 and were non-repeating. The edge-to-edge separation between these equal-sized letters was always fixed at one letter width. Here the letter size and spacing covaried, which should have not affected crowding measurements because crowding is known to be limited by interletter spacing, not stimulus size (Coates, Chin, & Chung, 2013; Song, Levi, & Pelli, 2014). An advantage of this method is

that it avoids overlap between targets and flankers, which was important because the whole-report task in our study required reporting of all three letters over the full range of crowding.

Procedures

Brief presentation of the trigram (200 ms) was initiated by the observer by a key press after achieving

proper fixation. The observer's task was to report the central letter (partial report) or all three letters (whole report) in order from left to right (i.e., from the inner to the outer flanker) with number keys 0–9, each corresponding to one Sloan letter. A printout of the 10 Sloan letters and their corresponding number keys was always placed in front of the keyboard for reference. Auditory error feedback was given when the central letter was wrongly reported in the partial-report paradigm and after all letters were reported in the whole-report paradigm. Observers were informed that there were no repeating letters, so that they should avoid reporting the same letter more than once in a whole-report trial. Trials were excluded from data analysis if eye position deviated from the fixation point by more than 1.5° before stimuli offset. Those trials accounted for $4.9\% \pm 2.5\%$ of total trials.

We used the method of constant stimuli with six letter sizes to measure psychometric functions and estimate letter recognition thresholds. These letter sizes were determined for each observer in a 30-min initial practice session: The observer first practiced the partial report with stimulus sizes initially set at 25, 40, 55, 70, 85, and 100 arcmin. These sizes were the average stimulus sizes used by Zhang et al. (2012). Each stimulus was practiced for 10 trials. The stimulus sizes were then adjusted if necessary for each observer, so that the psychometric function covered a sufficient range of correct rates.

Experiments 1 (except for the control condition, see later) and 2 each consisted of seven daily sessions: one pretraining session, five training sessions, and one posttraining session. Each session took approximately 1.5 hr to complete. The pre- and posttraining sessions each consisted of three experimental conditions: single-letter identification, trigram identification partial report, and trigram identification whole report. Each condition consisted of five blocks of trials. Each block consisted of 60 trials, 10 trials per letter size. A total of 15 blocks were run following a permuted table. Each partial-report (Experiment 1) or whole-report (Experiment 2) training session consisted of 15 blocks of trials, 60 trials per block (10 trials per letter size), for a total of 900 trials. Six and eight observers participated in Experiments 1 and 2, respectively.

The control condition in Experiment 1 consisted of five daily sessions: one pretraining session, three training sessions, and one posttraining session. The pre- and posttraining sessions each consisted of four experimental conditions: single-letter identification, trigram identification partial and whole reports at the trained location, and trigram identification partial report at an untrained mirror location in the contralateral hemifield. Each condition except for single-letter identification consisted of four blocks of trials. Each block consisted of 120 trials, and each letter size was

tested in a mini-block of 20 trials, in a sequence from large to small size. The single-letter condition consisted of two blocks of trials, and each letter size was tested in a mini-block of 25 trials, also from large to small. A total of 14 blocks were run following a permuted table. In three training sessions, the observer practiced trigram identification partial report at a single stimulus size that corresponded to a 50% target report rate in the pretraining session. Each training session consisted of 10 blocks of trials, 80 trials per block. Six observers participated in this control experiment.

The psychometric functions were fitted with a Weibull function $P = 1 - (1 - \gamma)e^{-(x/th)^\beta}$, where P is the percentage correct, γ is the guessing rate (0.1 with partial report and 0.3 with whole report), x is the letter size, β is the slope of the psychometric function, and th is the threshold letter size at a 66.9% correct rate in the partial-report paradigm and a 74.2% correct rate in the whole-report paradigm. In the whole-report paradigm, reporting from left to right (inner flanker to outer flanker), the chance rates of the central and outer reports are conditional rates that are affected by the inner report rates. The 0.3 overall chance rate comes from the sum of the inner-report chance rate: $1/10$; the central-report chance rate: $(1 - 1/10)(1/9)$; and the outer-report chance rate: $(1 - 1/10)(1 - 1/9)(1/8)$.

Results

Experiment 1: The impact of partial-report training on crowding

We constructed individual psychometric functions (Figure 2b) for single-letter recognition, partial report, and whole report before and after training. Partial-report training reduced letter recognition thresholds (estimated from the Weibull fits, Figure 2c) by $33.6\% \pm 3.8\%$, from 83.5 ± 8.6 arcmin to 55.0 ± 5.6 arcmin ($p = 0.002$, one-tailed paired t test here and in later analyses unless specified). There was no significant change of the slope of the partial-report psychometric function ($p = 0.35$). In addition, training had no significant impact on single-letter recognition thresholds (23.1 ± 1.7 vs. 21.8 ± 0.6 , $p = 0.31$). We quantified crowding effects with a crowding index, based on the ratio of the flanked central-letter threshold to the single-letter threshold. The crowding index decreased by 25.6% ($p = 0.012$, Figure 2d), suggesting significantly reduced crowding after training.

We have previously shown that some partial-report errors are actually target misplacement errors (when the central letter is actually correctly recognized but wrongly perceived in a flanker position; Zhang et al., 2012). If these target misplacement errors are counted

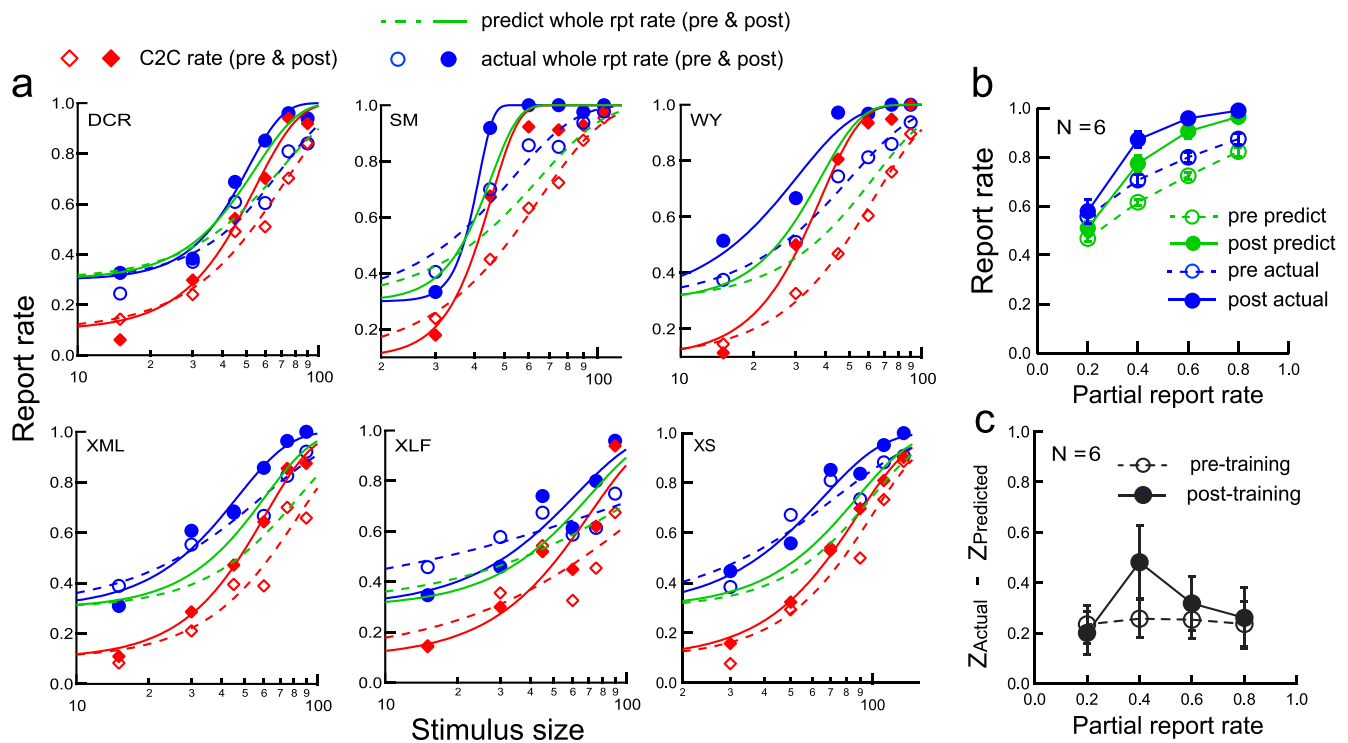


Figure 3. The effects of partial-report training on target recognition and misplacement errors. (a) Individual psychometric functions for C2C rates, predicted whole-report rates, and actual whole-report rates. Empty and solid symbols represent pre- and posttraining data, respectively. Dashed and solid red curves are Weibull fittings. (b) Pre- and posttraining predicted and actual whole-report rates at sizes corresponding to 0.2, 0.4, 0.6, and 0.8 partial-report rates in pretest. (c) Pre- and posttraining target misplacement errors in differences of z scores.

as correct target recognition, as shown in whole-report data, the flanked letter recognition pretraining threshold is lower, at 68.2 ± 10.5 arcmin pretraining, and is further reduced to 47.9 ± 5.3 arcmin posttraining ($p = 0.018$, Figure 2c). The corresponding crowding index is reduced by 19.6% ($p = 0.046$, Figure 2d).

The impact of partial-report training on target recognition and misplacement errors

In the whole-report paradigm, if the central-letter target is reported to a flanker position by chance, the whole-report rate of the central letter (regardless of the reported position) can be predicted from the rate of reporting the central letter in the central location (C2C rate). Specifically, the psychometric function of predicted whole-report rate can be transformed from the C2C Weibull fitting function with the same threshold and slope but with a chance level of 0.3 instead of 0.1. The C2C rates, predicted whole-report rates, and actual whole-report rates are presented in Figure 3a. Moreover, to make data comparable among observers, we further estimated predicted and actual whole-report rates at four letter sizes, where the pretraining partial report rates were 0.2, 0.4, 0.6, and 0.8, respectively, on the basis of Weibull fittings (Figure 3b).

To assess the impact of training on target recognition, we took the whole-report rate as the actual target recognition rate (ignoring the target's reported position). The pre- and posttraining actual whole-report rates were compared to a repeated-measures ANOVA with Training (pre- vs. posttest) and Size (four letter sizes corresponding to 0.2, 0.4, 0.6, and 0.8 partial-report rates in pretest) as within-subject factors. The results show a significant main effect of Training on actual whole-report rates, $F(1, 5) = 59.42$, $p = 0.001$, confirming improved target letter recognition with training. There was also a significant interaction between Training and Size, $F(3, 15) = 4.08$, $p = 0.026$. Pair-wise comparisons indicated that this interaction was mainly due to the nonsignificant change of report rate at the smallest letter size ($p = 0.55$).

To assess changes in the proportion of target misplacement errors, we transformed the actual and predicted whole-report rates to z scores and took the z-score differences as target misplacement errors (Figure 3c). A repeated-measures ANOVA with Training (pre- vs. posttest) and Size (0.2, 0.4, 0.6, and 0.8) as within-subject factors showed a nonsignificant main effect of Training, $F(1, 5) = 1.91$, $p = 0.23$, indicating that partial-report training had no significant impact on target misplacement errors. There was no significant

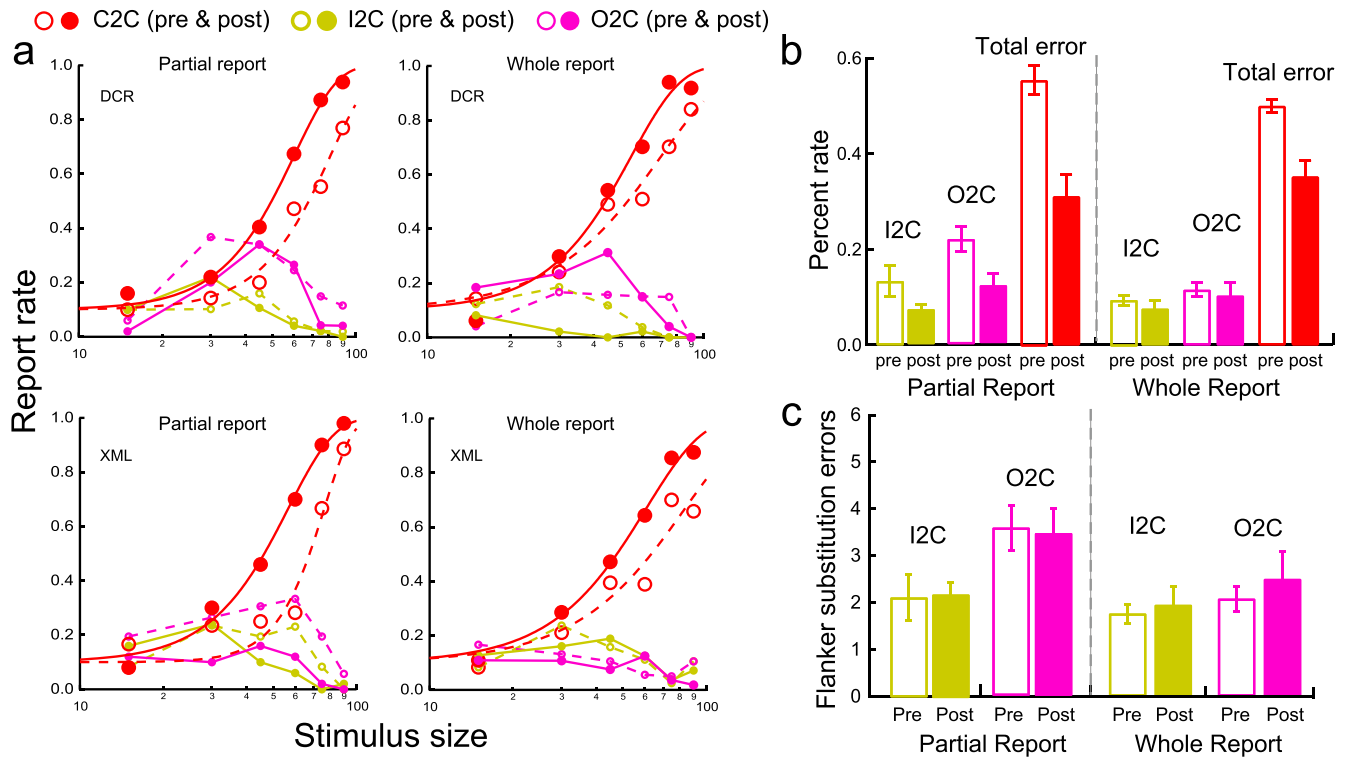


Figure 4. The effects of partial-report training on flanker substitution errors. (a) Examples of two observers showing C2C, I2C, and O2C rates at various stimulus sizes in partial report (left panels) and whole report (right panels). Empty and solid symbols represent data of pre- and posttraining, respectively. Dashed and solid red curves for pre- and posttraining C2C data are Weibull fittings. (b) Group means of pre- and posttraining I2C and O2C error rates, as well as total error rates ($1 - \text{C2C}$) in partial and whole reports. (c) Group means of normalized pre- and posttraining I2C and O2C error rates in partial and whole report.

main effect of Size, $F(3, 15) = 1.39$, $p = 0.28$, and no significant interaction between Training and Size, $F(3, 15) = 0.77$, $p = 0.53$.

The impact of partial-report training on flanker substitution errors

When failing to recognize the central letter at the central location, the observer is more likely to report a flanker letter than a no-show letter. Figure 4a presents pre- and posttraining inner-flanker (I2C) and outer-flanker (O2C) substitution errors in partial and whole reports in two of the six observers as examples. To analyze the impact of partial-report training on flanker substitution, we obtained normalized I2C and O2C errors for each observer: We first averaged I2C and O2C error rates, respectively, at stimulus sizes that resulted in 0.2 to 0.8 partial-report rates for each observer. The average I2C and O2C rates were then divided by the total error rates ($1 - \text{C2C}$) and then by the chance level ($1/9$). The group means of I2C and O2C error rates are shown in Figure 4b, and their normalized values are shown in Figure 4c.

A repeated-measures ANOVA with Training (pre- vs. posttest), Error (I2C vs. O2C), and Report (partial vs. whole report) as within-subject factors showed no

significant main effects of Training, $F(1, 5) = 0.24$, $p = 0.65$, or Error, $F(1, 5) = 4.64$, $p = 0.08$. However, there was a significant main effect of Report, $F(1, 5) = 17.33$, $p = 0.009$, due to more flanker substitution errors in partial report, which as pair-wise comparisons indicated was contributed by the difference of O2C errors, $F(1, 5) = 19.24$, $p = 0.007$, but not the difference of I2C errors, $F(1, 5) = 0.29$, $p = 0.61$. There were no significant interactions between Training and Report, $F(1, 5) = 1.00$, $p = 0.36$, or Training and Error, $F(1, 5) = 0.001$, $p = 0.98$. These results suggest that although partial-report training reduced crowding, it had no significant impact on flanker substitution errors. In addition, more flanker substitution errors, mainly O2C errors, were associated with partial report than with whole report.

A control condition: Partial-report training at a single stimulus size

We ran a control condition to address a number of issues. First, the flanker size and locations varied every trial during earlier training. This size and positional noise may have influenced the training effects on target misplacement and flanker substitution errors. Second,

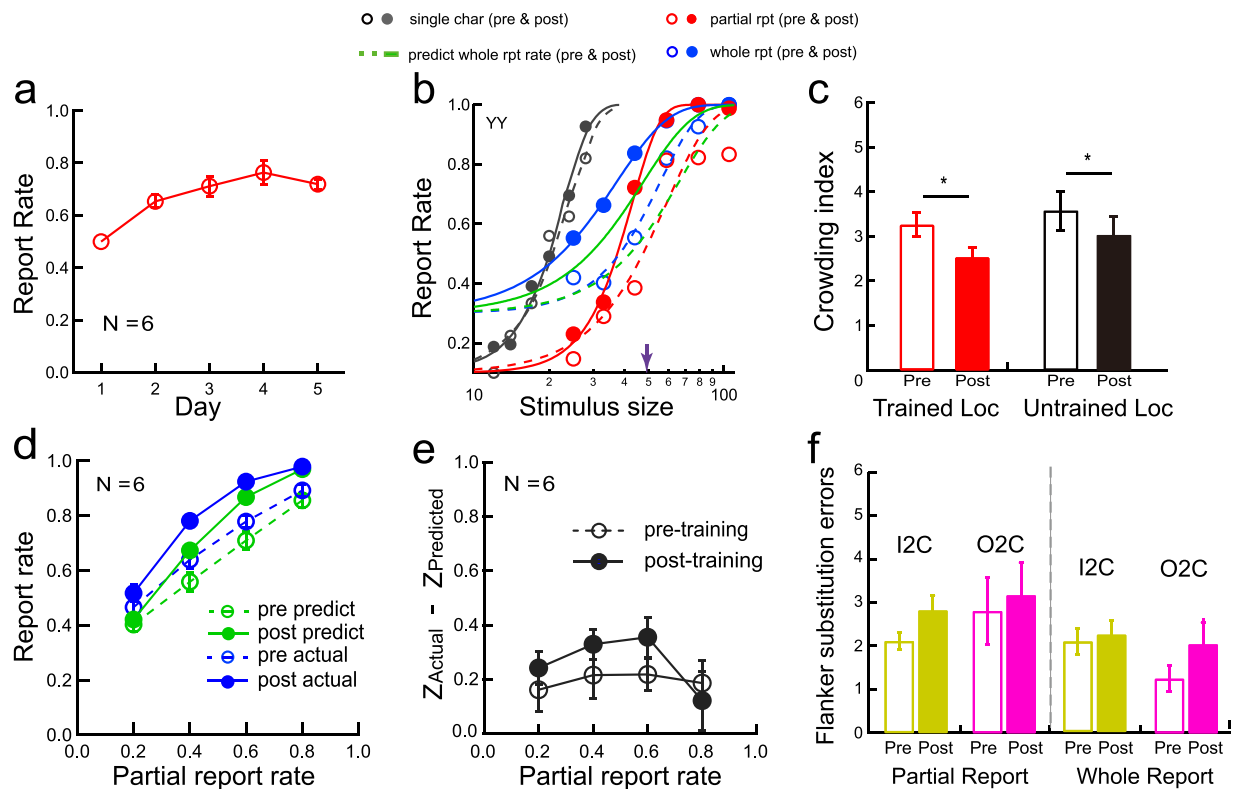


Figure 5. Control condition with partial-report training at a single letter size. (a) Partial-report rate for the trained letter size as a function of partial-report training. (b) One observer's pre- and posttraining psychometric functions for single-letter recognition rates, partial-report rates, predicted whole-report rates, and actual whole-report rates. Empty and solid symbols represent pre- and posttraining data, respectively. Dashed and solid curves are Weibull fittings. A purple arrow indicates the letter size that resulted in 0.5 partial-report rate in pretest. (c) Pre- and posttraining crowding indices in partial report at trained and untrained locations. (d) Pre- and posttraining predicted and actual whole-report rates at letter sizes corresponding to 0.2, 0.4, 0.6, and 0.8 partial-report rates in pretest. (e) Pre- and posttraining target misplacement errors in differences of z scores. (f) Normalized flanker substitution errors pre- and posttraining in partial and whole report.

there may not have been sufficient training for each specific stimulus condition. Third, each letter size was tested 50 times in pre- and posttraining measurements, which might not be sufficient to build accurate psychometric functions and provide precise estimates of the amount of learning. In the new control condition, training was performed at a constant letter size that led to a 50% partial-report rate in a pretraining session, during which each letter size was run 80 times. The pre- and posttraining psychometric functions were measured with stimulus size changed from large to small in mini-blocks to further reduce stimulus uncertainty (see Methods). In addition, the crowding effect was remeasured at an untrained symmetrical location in the contralateral hemifield to examine the location specificity of learning.

Partial-report training at a fixed stimulus size produced similar results to those from earlier training (Figures 2–4). The partial-report rate at the trained stimulus size was improved by $30.2\% \pm 1.8\%$ (Figure 5a), and the learning benefited the entire psychometric functions of both partial and whole reports (Figure 5b).

The central-letter recognition threshold was reduced by $25.3\% \pm 2.6\%$ ($p < 0.001$) in partial report and by $26.0\% \pm 6.1\%$ ($p = 0.010$) in whole report, but the single-letter recognition threshold was not significantly changed ($p = 0.24$). The central-letter recognition threshold in partial report at an untrained contralateral location was also reduced by $18.1\% \pm 5.0\%$ ($p = 0.007$). The crowding index in partial report was reduced significantly, by $22.6\% \pm 2.3\%$ at the trained location and $14.8\% \pm 5.6\%$ at the untrained location (Figure 5c). The untrained/trained ratio of crowding-index change was 0.58 ± 0.20 , indicating partial location specificity of learning.

Changes in target recognition rates

The target recognition rates, or whole-report rates, at letter sizes corresponding to 0.2, 0.4, 0.6, and 0.8 partial-report rates in pretest were improved (Figure 5d). A repeated-measures ANOVA with Training (pre- vs. posttest) and Size (0.2, 0.4, 0.6, and 0.8) as within-subject factors showed a significant main effect of

Training, $F(1, 5) = 23.14$, $p = 0.005$, confirming improved target recognition with training. The ANOVA also showed a significant interaction of Training and Size, $F(3, 15) = 6.67$, $p = 0.004$. Pair-wise comparisons indicate that this interaction was caused by a nonsignificant improvement at 0.2 letter size ($p = 0.12$).

Changes of target misplacement errors

The actual and predicted whole-report rates were transformed to z scores, and the z-score differences were taken as target misplacement errors (Figure 5e). A repeated-measures ANOVA with Training (pre- vs. posttest) and Size (0.2, 0.4, 0.6, and 0.8) as within-subject factors showed a nonsignificant main effect of Training, $F(1, 5) = 1.57$, $p = 0.27$, indicating that partial-report training had no significant impact on target misplacement errors. There was no significant main effect of Size, $F(3, 15) = 1.12$, $p = 0.37$, and no significant interaction between Training and Size, $F(3, 15) = 1.98$, $p = 0.16$.

Changes of flanker substitution errors

A repeated-measures ANOVA with Training (pre- vs. posttest), Error (I2C vs. O2C), and Report (partial vs. whole report) as within-subject factors showed a significant main effect of Training on normalized flanker substitution errors, $F(1, 5) = 10.58$, $p = 0.023$ (Figure 5f). However, the flanker substitution rates were increased rather than decreased. The ANOVA also showed a significant main effect of Report, $F(1, 5) = 11.66$, $p = 0.019$, due to more O2C errors in partial report, consistent with the earlier data (Figure 4c). Thus partial-report training at a single letter size did not reduce flanker substitution errors either.

Experiment 2: The impact of whole-report training on crowding

Partial-report training in Experiment 1 was focused on recognition of the central target, which might not be an effective way to reduce errors of reporting the target to a flanker position or a flanker to the target position. Therefore, in Experiment 2 we investigated the impact of whole-report training on crowding and response errors (Figure 6a).

Whole-report training reduced the central-letter recognition threshold by $19.3\% \pm 4.1\%$ ($p = 0.002$) in partial report and by $28.3\% \pm 5.6\%$ ($p = 0.001$) in whole report (Figure 6b), but training did not significantly change the single-letter recognition threshold ($p = 0.06$). The crowding index was reduced

by $15.8\% \pm 5.0\%$ in partial report and $25.7\% \pm 5.2\%$ in whole report (Figure 6c).

The effects of whole-report training on target recognition and target misplacement errors

As in Experiment 1, the pre- and posttraining target recognition rates, or whole-report rates, were estimated at four letter sizes corresponding to 0.2, 0.4, 0.6, and 0.8 partial-report rates in pretest on the basis of Weibull fittings (Figure 7a, b). A repeated-measures ANOVA with Training (pre- vs. posttraining) and Size (0.2, 0.4, 0.6, and 0.8) as within-subject factors showed significant main effects of Training, $F(1, 7) = 37.11$, $p < 0.001$, and Size, $F(3, 21) = 96.14$, $p < 0.001$, as well as a significant interaction of Training and Size, $F(3, 21) = 3.74$, $p = 0.027$, due to more improvement of recognition rates at 0.4 and 0.6 letter sizes. These results indicated that the whole-report training also significantly improved target recognition, especially at medium stimulus sizes.

To assess the training impact on target misplacement errors, again the actual and predicted whole-report rates were transformed to z scores, and the z-score differences were taken as target misplacement errors (Figure 7c). A repeated-measures ANOVA with Training (pre- vs. posttest) and Size (0.2, 0.4, 0.6, and 0.8) as within-subject factors showed a significant main effect of Training, $F(1, 7) = 10.47$, $p = 0.014$, but the target misplacement errors were increased rather than decreased. Thus there is no evidence for reduced target misplacement errors by whole-report training either. In addition, there was no significant main effect of Size, $F(3, 21) = 0.29$, $p = 0.83$, and no significant interaction between Training and Size, $F(3, 21) = 0.18$, $p = 0.91$.

The effect of whole-report training on flanker substitution errors

Again we averaged I2C and O2C error rates, respectively, at stimulus sizes that resulted in 0.2 to 0.8 partial-report rates for each observer. The average I2C and O2C rates were divided by the total error rates ($1 - C2C$) and then by the chance level ($1/9$), to obtain normalized I2C and O2C errors for each observer. A repeated-measures ANOVA with Training (pre- vs. posttest), Error (I2C vs. O2C), and Report (partial vs. whole report) as within-subject factors showed no significant main effect of Training, $F(1, 7) = 0.14$, $p = 0.72$ (Figure 8c). However, there was a significant main effect of Error, $F(1, 7) = 25.34$, $p = 0.002$, due to more O2C errors than I2C errors, as well as a significant main effect of Report, $F(1, 7) = 101.81$, $p = 0.00$, due to more errors in partial report than in whole report. Pair-wise comparisons showed significantly higher I2C ($p = 0.007$) and O2C ($p < 0.001$) errors in partial report

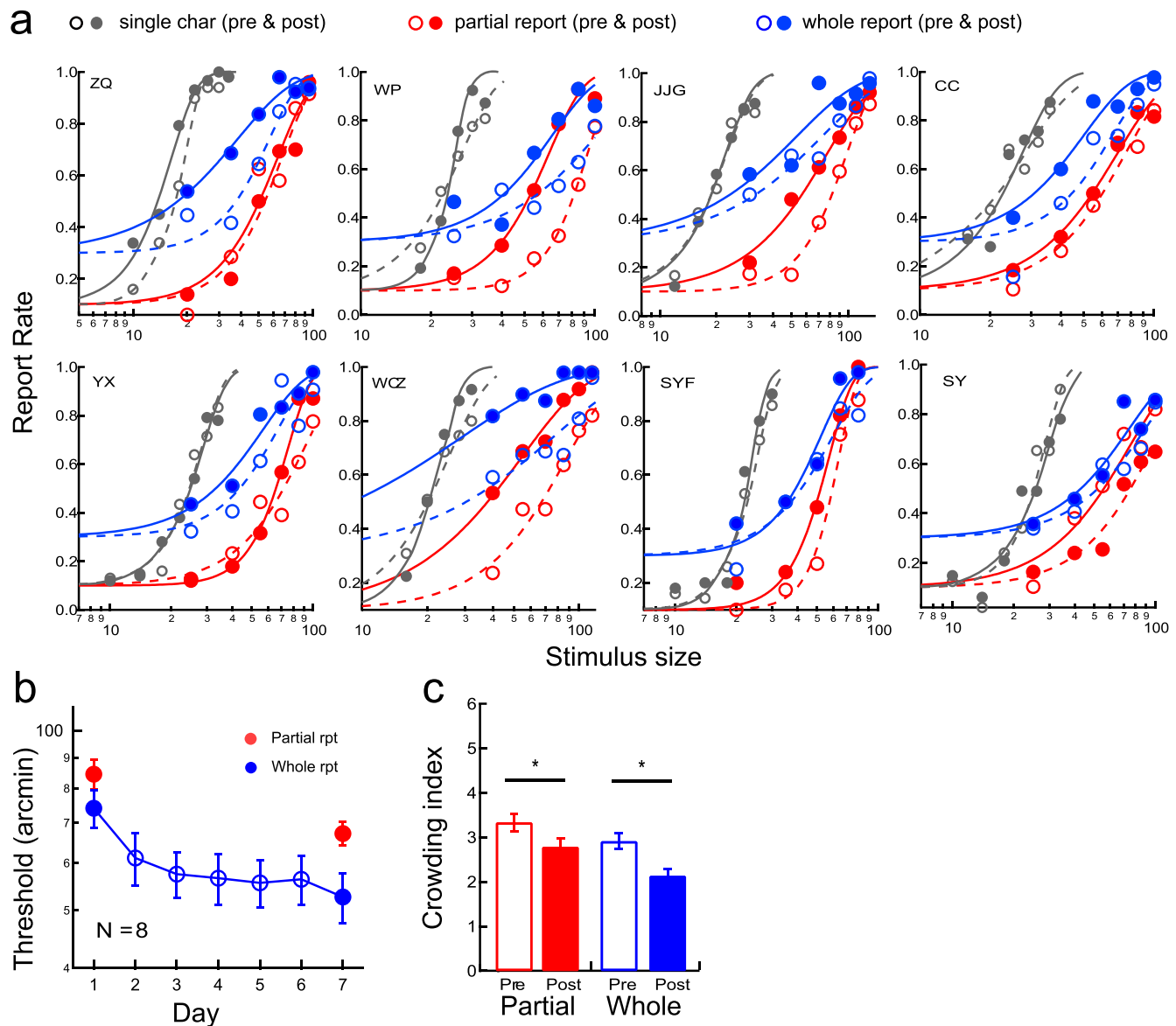


Figure 6. The effects of whole-report training on crowding. (a) Individual pre- and posttraining psychometric functions for single-letter recognition, partial report, and whole report. (b) Letter recognition thresholds estimated with whole-report data through the course of whole-report training, as well as pre- and posttraining thresholds estimated with partial-report data. (c) Pre- and posttraining crowding indices in partial and whole report.

than in whole report. There were no significant interactions between Training and Report, $F(1, 7) = 0.16$, $p = 0.71$, or Training and Error, $F(1, 7) = 2.06$, $p = 0.20$. Therefore, like partial-report training, whole-report training had no significant impact on normalized flanker substitution errors.

Discussion

By investigating training-induced changes of various response errors associated with crowding, we are able to have a detailed look into perceptual learning that

reduces crowding in peripheral letter recognition. Our results show that both partial- and whole-report training improve recognition of the flanked letter target. Critically, single-letter recognition is unaffected, suggesting that training indeed reduces the crowding effect on letter recognition rather than letter recognition per se. On the other hand, partial- and whole-report training do not reduce target misplacement errors and flanker substitution errors.

Strasburger (2005) has suggested that flanker substitution, instead of being a cause of crowding, could reflect the observer's bias to report a more visible flanker when failing to recognize the central target. In this case, flanker substitution is a by-product of crowding. Our training

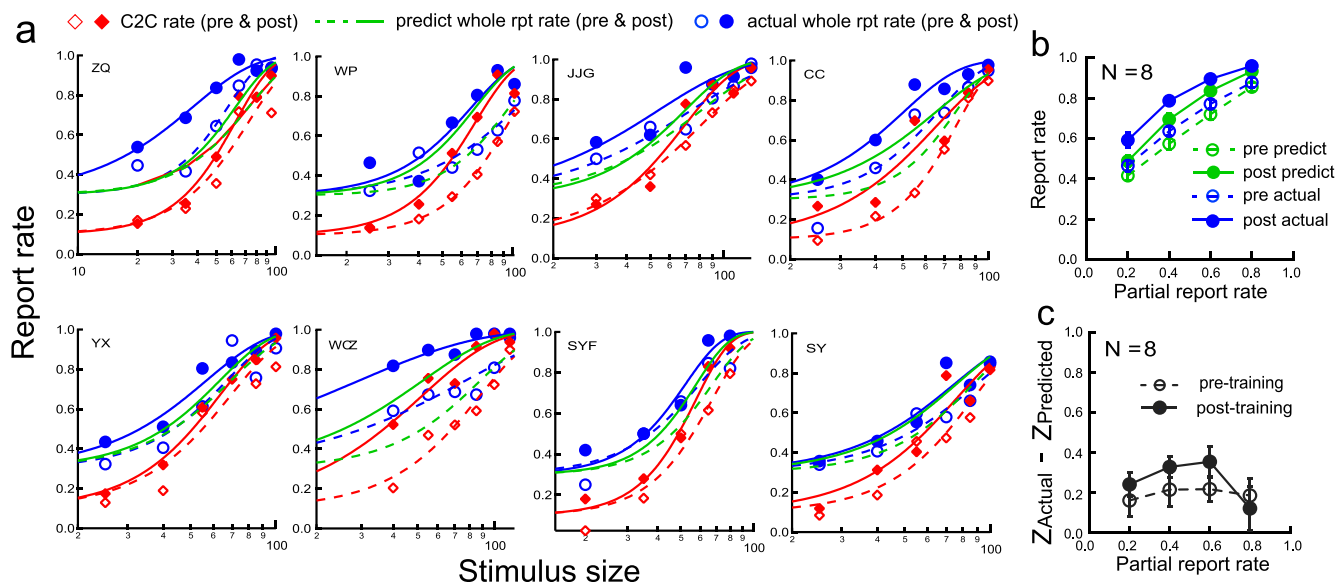


Figure 7. The effects of whole-report training on target recognition and target misplacement errors. (a) Individual psychometric functions for C2C rates, predicted whole-report rates, and actual whole-report rates. Empty and solid symbols represent pre- and posttraining data, respectively. Dashed and solid red curves are Weibull fittings. (b) Pre- and posttraining predicted and actual whole-report rates at sizes corresponding to 0.2, 0.4, 0.6, and 0.8 partial-report rates in pretest. (c) Pre- and posttraining target misplacement errors in differences of z scores.

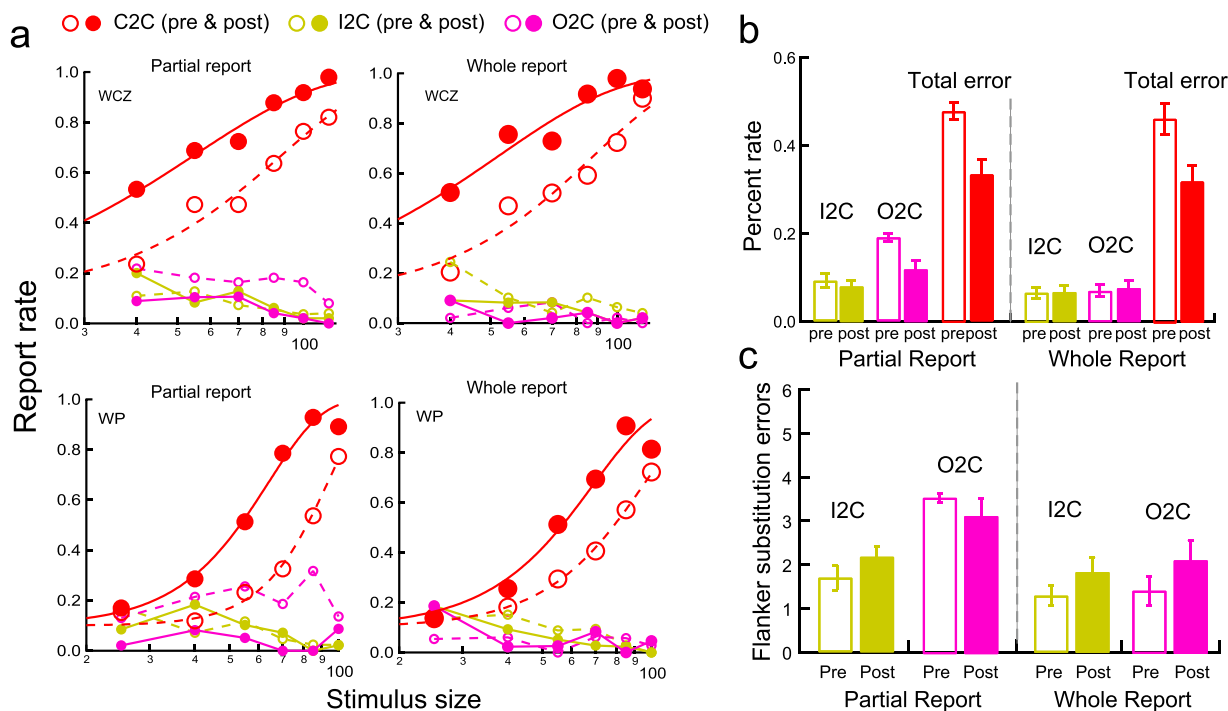


Figure 8. The effects of whole-report training on flanker substitution errors. (a) Examples of two observers showing C2C, I2C, and O2C rates at various stimulus sizes in partial report (left panels) and whole report (right panels). Empty and solid symbols represent data pre- and posttraining, respectively. Dashed and solid red curves for pre- and posttraining C2C data are Weibull fittings. (b) Group means of pre- and posttraining I2C and O2C error rates, as well as total error rates (1 – C2C) in partial and whole report. (c) Group means of normalized pre- and posttraining I2C and O2C error rates in partial and whole report.

results dissociate crowding and flanker substitution errors, consistent with the response-bias explanation. In addition, there are fewer flanker substitution errors in whole-report tasks than in partial-report tasks, and this difference can be mainly attributed to reduced O2C errors (Figures 4, 5, and 8). These results are also consistent with the response-bias explanation, because the outer flanker is the most visible stimulus in a trigram (Bouma, 1970). Recently Y. He et al. (2013) reported that training reduces not only recognition errors of the flanked target but also mislocation errors that include both target misplacement and flanker substitution errors in our terms. However, their mislocation errors are not normalized by the corresponding recognition errors. We analyzed recognition and mislocation error data kindly provided by those authors, and found that indeed the mislocation errors are unchanged if normalized. The flanker substitution errors may result from interactions between uneven visibilities of target and flankers in a trigram and intrinsic positional uncertainty in the visual periphery, as documented by many studies (Klein & Levi, 1987; Levi & Klein, 1986; Levi, Klein, & Yap, 1987; Michel & Geisler, 2011). The impact of these interactions is probably more prominent in a forced-choice report, as is the case in most crowding studies.

On the other hand, target misplacement errors are more likely perceptual and/or memory errors, rather than report errors. In the whole-report paradigm, the observer needs to report all three letters in order. It is less likely for the observer to wrongly report a target to a more visible flanker position. Besides perceptual errors, the positions of the target and a flanker could be exchanged in working memory. These errors may also be caused by intrinsic positional uncertainty in the visual periphery (Klein & Levi, 1987; Levi et al., 1987; Levi & Klein, 1986; Michel & Geisler, 2011), which are not directly targeted by our training including whole-report training.

Although positional uncertainty may play a role in both flanker substitution errors and target misplacement errors, it may affect the two types of errors differently. For example, in a typical (partial-report) crowding experiment, positional uncertainty may be responsible for target misplacement errors that are present as part of the report errors. When a wrong report does occur, positional uncertainty may also interact with a more visible flanker to produce a flanker substitution error.

A combination of improved target recognition and unchanged positional errors in our training data provides a unique opportunity to examine some influential explanations of crowding. If training eases crowding via improving the resolution of spatial attention (S. He et al., 1996; Intriligator & Cavanagh, 2001) and its precise allocation (Strasburger, 2005), positional errors including target misplacement errors and flanker substitution

errors are expected to be reduced. However, this prediction is not supported by our results. On the other hand, if training reduces unwanted feature integration between target and flankers (Greenwood et al., 2009; Levi et al., 2002; Pelli et al., 2004), target misplacement and flanker substitution as positional errors may not be affected, since feature integration errors and positional errors are not necessarily related to each other. Our results are thus more consistent with the feature-integration explanation of crowding.

The partial location specificity of perceptual learning in the control condition (Figure 5) confirms the involvement of perceptual learning in training, which is consistent with a previous report (Yeotikar, Khoo, Asper, & Suttle, 2013). However, what we can learn from this partial location specificity requires further study. Many visual-learning tasks are known to be at least partially specific to the trained location, but we—and, more recently other labs too—have demonstrated that location-specific perceptual learning can often be rendered completely transferrable to other untrained locations with double training (Hung & Seitz, 2014; Mastropasqua, Galliussi, Pascucci, & Turatto, 2015; Wang, Cong, & Yu, 2013; Wang, Zhang, Klein, Levi, & Yu, 2012, 2014; Xiao et al., 2008). We are currently investigating whether double training can be equally effective in eliminating location specificity in crowding-related perceptual learning, with the intention of using the transfer effects to infer the mechanisms underlying crowding and its learning.

Keywords: crowding, perceptual learning, target misplacement, flanker substitution

Acknowledgments

We thank Dennis Levi for helping improve the writing of the manuscript. This research was supported by Natural Science Foundation of China Grants 31470975 and 31230030.

Commercial relationships: none.

Corresponding author: Jun-Yun Zhang.

Email: zhangjunyun@pku.edu.cn.

Address: Department of Psychology and Peking–Tsinghua Center for Life Sciences, Peking University, Beijing, China.

References

- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226(5241), 177–178.
- Chung, S. T. (2007). Learning to identify crowded

- letters: Does it improve reading speed? *Vision Research*, 47(25), 3150–3159.
- Chung, S. T., Li, R. W., & Levi, D. M. (2012). Learning to identify near-acuity letters, either with or without flankers, results in improved letter size and spacing limits in adults with amblyopia. *PLoS One*, 7(4), e35829, doi:10.1371/journal.pone.0035829.
- Coates, D. R., Chin, J. M., & Chung, S. T. (2013). Factors affecting crowded acuity: Eccentricity and contrast. *Optometry & Vision Science*, 90(7), 628–638, doi:10.1097/OPX.0b013e31829908a4.
- Flom, M. C., Heath, G. G., & Takahashi, E. (1963, Nov 15). Contour interaction and visual resolution: Contralateral effects. *Science*, 142, 979–980.
- Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences, USA*, 106(31), 13130–13135.
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, 383(6598), 334–337.
- He, Y., Legge, G. E., & Yu, D. (2013). Sensory and cognitive influences on the training-related improvement of reading speed in peripheral vision. *Journal of Vision*, 13(7):14, 1–14, doi:10.1167/13.7.14. [PubMed] [Article]
- Huckauf, A., & Heller, D. (2002). What various kinds of errors tell us about lateral masking effect. *Visual Cognition*, 9(7), 889–910.
- Huckauf, A., & Nazir, T. A. (2007). How odgcrnwi becomes crowding: Stimulus-specific learning reduces crowding. *Journal of Vision*, 7(2):18, 1–12, doi:10.1167/7.2.18. [PubMed] [Article]
- Hung, S. C., & Seitz, A. R. (2014). Prolonged training at threshold promotes robust retinotopic specificity in perceptual learning. *The Journal of Neuroscience*, 34(25), 8423–8431, doi:10.1523/JNEUROSCI.0745-14.2014.
- Hussain, Z., Webb, B. S., Astle, A. T., & McGraw, P. V. (2012). Perceptual learning reduces crowding in amblyopia and in the normal periphery. *The Journal of Neuroscience*, 32(2), 474–480, doi:10.1523/JNEUROSCI.3845-11.2012.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, 43(3), 171–216.
- Klein, S. A., & Levi, D. M. (1987). Position sense of the peripheral retina. *Journal of the Optical Society of America A*, 4(8), 1543–1553, doi:10.1364/Josaa.4.001543.
- Krumhansl, C. L. (1977). Naming and locating simultaneously and sequentially presented letters. *Perception & Psychophysics*, 22, 293–302.
- Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, 48(5), 635–654.
- Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision*, 2(2):3, 167–177, doi:10.1167/2.2.3. [PubMed] [Article]
- Levi, D. M., & Klein, S. A. (1986). Sampling in spatial vision. *Nature*, 320(6060), 360–362, doi:10.1038/320360a0.
- Levi, D. M., Klein, S. A., & Yap, Y. L. (1987). Positional uncertainty in peripheral and amblyopic vision. *Vision Research*, 27(4), 581–597, doi:10.1016/0042-6989(87)90044-7.
- Mastropasqua, T., Galliussi, J., Pascucci, D., & Turatto, M. (2015). Location transfer of perceptual learning: Passive stimulation and double training. *Vision Research*, 108, 93–102.
- Michel, M., & Geisler, W. S. (2011). Intrinsic position uncertainty explains detection and localization performance in peripheral vision. *Journal of Vision*, 11(1):18, 1–18, doi:10.1167/11.1.18. [PubMed] [Article]
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, 4(12), 1136–1169, doi:10.1167/4.12.12. [PubMed] [Article]
- Song, S., Levi, D. M., & Pelli, D. G. (2014). A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. *Journal of Vision*, 14(5):3, 1–37, doi:10.1167/14.5.3.
- Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, 5(11):8, 1024–1037, doi:10.1167/5.11.8. [PubMed] [Article]
- 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]
- Wang, R., Cong, L. J., & Yu, C. (2013). The classical TDT perceptual learning is mostly temporal learning. *Journal of Vision*, 13(5):9, 1–9, doi:10.1167/13.5.9. [PubMed] [Article]

- Wang, R., Zhang, J. Y., Klein, S. A., Levi, D. M., & Yu, C. (2012). Task relevancy and demand modulate double-training enabled transfer of perceptual learning. *Vision Research*, *61*, 33–38.
- Wang, R., Zhang, J. Y., Klein, S. A., Levi, D. M., & Yu, C. (2014). Vernier perceptual learning transfers to completely untrained retinal locations after double training: A “piggybacking” effect. *Journal of Vision*, *14*(13):12, 1–10, doi:10.1167/14.13.12. [PubMed] [Article]
- 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.
- Yeotikar, N. S., Khuu, S. K., Asper, L. J., & Suttle, C. M. (2013). Context and crowding in perceptual learning on a peripheral contrast discrimination task: Context-specificity in contrast learning. *PLoS One*, *8*(5), e63278, doi:10.1371/journal.pone.0063278.
- Zhang, J. Y., Zhang, G. L., Liu, L., & Yu, C. (2012). Whole report uncovers correctly identified but incorrectly placed target information under visual crowding. *Journal of Vision*, *12*(7):5, 1–11, doi:10.1167/12.7.5. [PubMed] [Article]