Linking crowding, visual span, and reading

Yingchen He

Department of Psychology, University of Minnesota, Twin Cities, MN, USA

Gordon E. Legge

Department of Psychology, University of Minnesota, Twin Cities, MN, USA

The visual span is hypothesized to be a sensory bottleneck on reading speed with crowding thought to be the major sensory factor limiting the size of the visual span. This proposed linkage between crowding, visual span, and reading speed is challenged by the finding that training to read crowded letters reduced crowding but did not improve reading speed (Chung, 2007). Here, we examined two properties of letter-recognition training that may influence the transfer to improved reading: the spatial arrangement of training stimuli and the presence of flankers. Three groups of nine young adults were trained with different configurations of letter stimuli at 10° in the lower visual field: a flanked-local group (flanked letters localized at one position), a flanked-distributed group (flanked letters distributed across different horizontal locations), and an isolated-distributed group (isolated and distributed letters). We found that distributed training, but not the presence of flankers, appears to be necessary for the training benefit to transfer to increased reading speed. Localized training may have biased attention to one specific, small area in the visual field, thereby failing to improve reading. We conclude that the visual span represents a sensory bottleneck on reading, but there may also be an attentional bottleneck. Reducing the impact of crowding can enlarge the visual span and can potentially facilitate reading, but not when adverse attentional bias is present. Our results clarify the association between crowding, visual span, and reading.

Introduction

Reading is important in many aspects of daily life, including access to books, medicine labels, or restaurant menus. People who have lost their central visual field from macular degeneration or other eye disease must rely on their peripheral vision to read. Reading in peripheral vision is slow and difficult, impairing quality of life. It is thus important to investigate what limits the reading performance in peripheral vision and how to improve it.

According to the visual span hypothesis, a sensory bottleneck on reading speed is the size of the visual span, which is the number of letters that can be recognized accurately without eye movements (Legge et al., 2007). The visual span hypothesis proposes that the shrinkage of the visual span accounts for slower reading speed in peripheral vision. By decomposing the errors made in visual span measurements, He, Legge, and Yu (2013) showed that visual crowding, the inability to recognize objects in clutter, is the major factor limiting the size of the visual span. Pelli et al. (2007) have provided direct evidence that the visual span for reading refers to the span of letters that are not crowded, and reading speed is proportional to this uncrowded span. It thus appears that crowding limits reading speed by limiting the size of the visual span.

If this linkage between visual span and reading speed is correct, reduced crowding should result in an enlarged visual span and improved reading speed. In support of this view, it was found that perceptual training on a trigram letter-recognition task in peripheral vision (Figure 1A) enlarged the visual span and improved reading speed with the major contributing factor being the reduction of crowding (He et al., 2013). But reduced crowding does not always result in improved reading. For example, reducing crowding by increasing letter spacing beyond standard spacing slows reading down (Chung, 2002; Yu, Cheung, Legge, & Chung, 2007). This finding does not necessarily argue against crowding as the major sensory limit on reading speed because enlarged spacing also resulted in larger eccentricity and degraded word form. Moreover, in this example, extra-wide letter spacing resulted in both a shrinkage of the visual span and a decrease in reading speed (Yu et al., 2007), confirming the link between visual span and reading.

But a further dissociation between crowding and reading was found by Chung (2007). She found that, after six days of training in identifying crowded letters
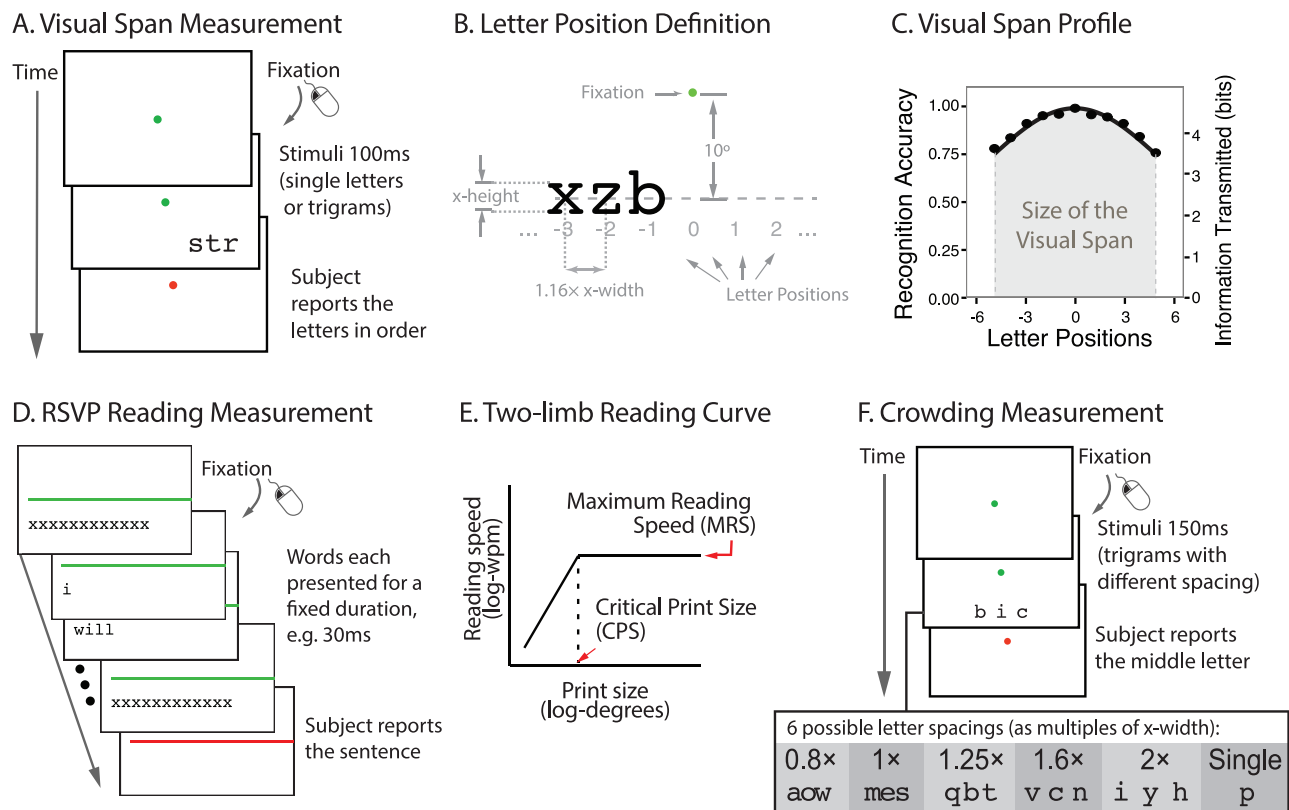
Figure 1. Diagrams of measurements (see Methods for detailed description of each task). (A) Visual span measurement. (B) Letter position definition used in visual span measurement. (C) Diagram of a visual span profile. (D) RSVP reading speed measurement. (E) Two-limb reading curve describing the influence of print size on reading speed. (F) Crowding measurement.

in peripheral vision (hereafter referred to as "uncrowd training"), the spatial extent of crowding (defined as the smallest target-to-flanker spacing yielding 50% recognition accuracy of the target letter) was reduced by 38%, but maximum reading speed barely changed. In a follow-up study, Chung and Truong (2013) showed that a similar training task could both reduce the spatial extent of crowding and enlarge the size of the visual span. These results argue against the proposed linkage between visual span, crowding, and reading speed.

Here, we examined two properties of the stimuli used in letter-recognition training that may influence the transfer to reading: the distributed or localized spatial arrangement of training stimuli and the presence or absence of flankers.

The first property is the spatial arrangement of training stimuli. In a typical trial of trigram visual span measurement (Figure 1A), a trigram appears briefly at a certain vertical eccentricity (say, 10° in the lower field) but at an unpredictable horizontal position. After the trigram disappears, the subject reports the identity of the three letters. From trial to trial, the horizontal location of the trigram varies, resembling different letter positions in a word relative to fixation. In a trial of "uncrowd" training as in Chung (2007), the

procedure is similar, but the trigram is always centered right below fixation. The trigram has narrower-than-standard letter spacing (0.8× *x*-width), and only the center (target) letter needs to be reported. As pointed out by Chung (2007), because the target letter always occupied the same location and the subjects might have learned to ignore the flanking letters, it is possible that improved letter-recognition performance only occurred at one letter position. For English reading, parallel processing of multiple letters is required to achieve fast reading. Letter-by-letter reading is extremely slow as observed in brain-lesioned adults with pure alexia (for a review, see Dehaene & Cohen, 2011). Therefore, Chung's (2007) "uncrowd" training may not transfer effectively to reading because of the restricted spatial range of the training stimuli. Our prediction is that when the uncrowd training is spatially distributed across locations, reading speed will show a greater improvement than when training occurs at a single spatial location.

A second property, pertinent to training peripheral vision to read, is whether crowding is present or not. That is, whether training trials use isolated letters or letters flanked by other letters. Excessive crowding is often associated with slow reading, not only in peripheral vision for normally sighted subjects (as we
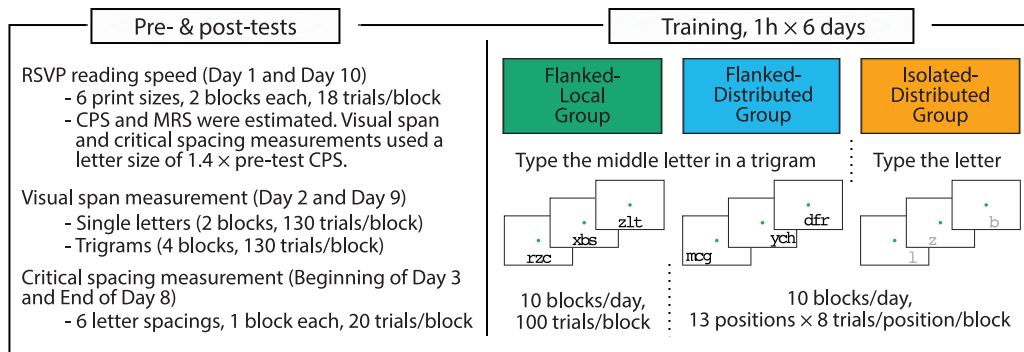
Figure 2. Experimental design. Our experiment consisted of three parts: pretest (2 days), training (6 days), and posttest (2 days). The subjects were assigned to three groups (flanked-local, flanked-distributed, and isolated-distributed). All three groups were tested with the same tasks in the pre- and posttests, but their training task differed. See text for detailed task descriptions.

discussed before), but also in central vision for people with amblyopia (Levi, Song, & Pelli, 2007) or dyslexia (Callens, Whitney, Tops, & Brysbaert, 2013; Martelli, Di Filippo, Spinelli, & Zoccolotti, 2009; Moll & Jones, 2013; but see Doron, Manassi, Herzog, & Ahissar, 2015). Given the association between crowding and reading speed, training without crowded stimuli may fail to reduce the effect of crowding and thus cannot improve reading. But if learning to identify unflanked letters can reduce crowding, reading speed should improve despite the fact that isolated letters only rarely occur in real-life text.

To study the influence of these two factors on training, we designed three training procedures for letter recognition (Figure 2): One procedure is a replication of Chung (2007) in which target letters are localized at 10° directly below fixation with closely spaced flankers on each side (flanked-local group). In the second procedure, target letters were located at 10° vertically below fixation but distributed across different horizontal locations, also with closely spaced flankers on each side (flanked-distributed group). The third procedure had isolated (no flankers), spatially distributed target letters (isolated-distributed). By evaluating the training effects on crowding, visual span, and reading speed, our goal was to investigate the necessary and sufficient conditions for transfer of training from letter recognition to the size of the visual span and reading speed.

Methods

Participants

Twelve male and 15 female college students were recruited from the University of Minnesota and assigned to three groups (described later). Participants all had normal or corrected-to-normal vision, and their binocular acuity (Lighthouse Near Acuity Chart, Lighthouse Low Vision Products, Long Island City, NY) and reading performance (MNREAD, Precision Vision, La Salle, IL) were tested before the experiment. MNREAD data were fitted for each individual using the exponential decay function described by Cheung, Kallie, Legge, and Cheong (2008). Subject information is summarized in Table 1. The protocol was approved by the institutional review board and was in compliance with the Declaration of Helsinki. All subjects gave informed consent prior to the experiment.

Stimuli and apparatus

The stimuli consisted of black lowercase letters on a white background (luminance 90 cd/m^2; Weber contrast = 99%) except for the training task of the isolated-distributed group for which stimuli were gray lowercase letters on a white background (Weber contrast ranges from 7% to 14%). All stimuli were viewed binocularly from 40 cm in a dark room. The letters were rendered in Courier font. Letter spacing in the reading task and visual span measurement was 1.16× x-width (standard spacing) but varied from 0.8× to 2× x-width in the crowding measurement.

The stimuli were generated and presented using MATLAB R2014b with Psychophysics Toolbox 3 (Brainard, 1997; Pelli, 1997). We used a NEC Multi-Sync CRT monitor (model FP2141SB-BK, NEC, Tokyo, Japan; refresh rate = 100 Hz; spatial resolution = 0.04°/pixel) controlled by a Mac Pro Quad-Core computer (model A1186, Apple Inc., Cupertino, CA). Viewing distance was maintained using a chin rest, and subject's fixation was monitored using a webcam. A previous study using a similar method has shown that saccades of 2° can be reliably detected by an experimenter (Cheong, Legge, Lawrence, Cheung, & Ruff, 2007). Trials were cancelled and replaced when fixation was not maintained.

| Group | Gender ratio (M:F) | Age | Visual acuity (logMAR) | MNREAD measurements | | |
				Reading acuity (logMAR)	Critical print size (logMAR)	Maximum reading speed (wpm)
Flanked-local	5:4	21.4 ± .7	−.07 ± .01	−.39 ± .03	.06 ± .05	221 ± 6
Flanked-distributed	3:6	22.7 ± 1.0	−.08 ± .01	−.4 ± .03	−.07 ± .04	234 ± 9
Isolated-distributed	4:5	21.3 ± .7	−.08 ± .01	−.4 ± .02	−.02 ± .03	213 ± 6
All groups	12:15	21.8 ± .4	−.08 ± .004	−.4 ± .02	.04 ± .02	223 ± 4

Table 1. Group characteristics (mean ± *SEM*).

Experimental design

The tasks we used are illustrated in Figure 1, and the procedure is explained in Figure 2. Our experiment consisted of three parts (Figure 2): pretest (2 days), training (6 days), and posttest (2 days). For 18 of the 27 subjects, the experiment took place on 10 consecutive days whereas for the other subjects scheduling arrangements meant that the experiment spanned 11 (four subjects), 12 (four subjects), or 15 (one subject) days. Because the effectiveness of visual perceptual learning on identifying crowded letters and enlarging the visual span is similar for daily, weekly, and biweekly training (Chung & Truong, 2013), the variations in scheduling likely had minimal effects on performance.

In the pre- and posttests, each subject's reading speed, visual span, and spatial extent of crowding were measured. The following sections describe each task as well as the training.

Reading measurement

Reading speed was measured on the first and last days of the experiment using rapid serial visual presentation (RSVP; Chung, Mansfield, & Legge, 1998; Forster, 1970; Rubin & Turano, 1992). For each RSVP trial, a sentence was randomly chosen from a pool of 847 sentences for testing. A subject never saw the same sentence twice. The average sentence length was 11 words, ranging from seven to 15 words. The word length averaged four letters, ranging from one to 12 letters. In an RSVP trial, a sentence was randomly chosen and presented word by word with the sentence preceded and followed by masks of "xxxxxxxxxx" (Figure 1D). Subjects were asked to fixate on a line without making vertical eye movements, but horizontal eye movements along the line were permitted. Words were presented 10° below the fixation line and were left-aligned with the left edge of the fixation line. Subjects read the sentences out loud and the experimenter recorded the number of correctly read words.

In each block of 18 trials, six different word exposure durations were tested in a random order (three times each). Depending on individual performance during practice, one of three duration sets could be chosen for a given block: 30, 53, 93, 164, 290, and 511 ms; 53, 93, 164, 290, 511, and 1000 ms; or 93, 164, 290, 511, 1000, and 2000 ms. The choice of set maximally ensures (a) greater than 80% word recognition accuracy for the longest duration in the set, (b) close to chance-level word recognition performance for the shortest duration in the set, and (c) good eye fixation for all durations (fixation becomes poorer as duration becomes longer). The resulting accuracy–duration curve was then fitted with a psychometric function, and the subject's reading speed (measured in words per minute [wpm]) was calculated using the exposure duration yielding 80% accuracy of word recognition (Chung et al., 1998).

Reading speed was measured in this way for six different print sizes: 0.56°, 0.79°, 1.12°, 1.59°, 2.26°, and 3.2° in *x*-height. In the pretest, these six print sizes were tested in a random order, one print size in a block, for the first six blocks and then tested in the reverse order for another six blocks. The posttest followed the same order as in the pretest. The resulting reading speed–print size curve was then fitted with a two-limb function on a log–log scale (Figure 1E) to extract the subject's maximum reading speed (MRS) and critical print size (CPS). The slope of the two limbs were constrained to 2.32 and 0 (following Chung et al., 1998, and Chung, 2007). This curve represents that reading speed remains constant at the MRS for larger print sizes but starts to decrease when print size becomes smaller than the CPS. After extracting the subject's pretest CPS, a print size of 1.4 × CPS was used for subsequent visual span measurement, crowding measurement, and training.

Visual span measurement

Visual span profiles were measured in the pre- and posttests (days 2 and 9) using a letter-recognition task as described in He et al. (2013). Stimuli were randomly chosen lowercase letters, either in isolation or arranged as trigrams (three adjacent letters). For most subjects, we measured the visual span profiles using both single letters and trigrams with some exceptions for the flanked-local group: Because this group of subjects was

recruited first while the measurement of single letters was added later in the data-collection process, five subjects in this group had no data on single-letter visual span, three subjects had the posttest data only, and only one subject had both the pre- and posttest data.

Figure 1A illustrates the basic procedure of a trial: A subject fixated on a dot and clicked the mouse to initiate a trial. An isolated letter or a trigram appeared for 100 ms in the lower visual field for the subject to identify. For a letter in a trigram to be correct, both its identity and location needed to be correct. Letters were presented in predefined slots as shown in Figure 1B. The slots were horizontally arranged on an imaginary line at 10° in the lower visual field. The slot on the fixation midline was labeled 0, and left and right slots were labeled with negative and positive numbers, respectively. The center-to-center spacing between adjacent slots was 1.16× x-width, corresponding to standard spacing in the Courier font. Because each individual used a different print size determined by their CPS for reading, the actual horizontal eccentricity of the slots varied across subjects. We analyzed the data using a "letter position" metric instead of degrees of visual angle because once print size exceeds the CPS for reading, the size of the visual span (in terms of number of recognizable letters) remains constant within a wide range of print sizes (Legge et al., 2007). By using the "letter position" metric, we were able to make direct comparisons between subjects.

In each block of trials, letters or trigrams were centered 10 times on each slot from −6 to 6, including 0 (the midline). This means that a total of 130 letters were presented during a single-letter block or 390 letters during a trigram block. In a trigram test, because slots ±6 and ±7 had fewer letters than the other slots, only data from slots −5 to 5 were used in further analyses.

To get the visual span profile, letter-recognition accuracy was plotted against letter position (Figure 1C). We converted letter-recognition accuracy to information transmitted in bits using the formula

$$\text{information transmitted in bits} = -0.036996 + 4.6761 \times \text{letter recognition accuracy},$$

where chance-level performance (about 3.8% correct) corresponds to zero bits of information, and 100% accuracy corresponds to about 4.7 bits. The number of bits was added for slots −5 to 5 to estimate the size of the visual span (Figure 1C).

Crowding measurement

The crowding task measured the spatial extent of crowding. It only took 10 min, and therefore, we performed the measurement immediately before the first training block on day 3 and after the last training block on day 8, similar to Chung (2007). In one trial (Figure 1F), a target letter appeared briefly (150 ms) on the screen either in isolation or with flanking letters on its left and right sides. The target letter was always placed at 10° in the lower visual field directly below the fixation point. The center-to-center spacing between the target and its flanking letters could be 0.8×, 1×, 1.25×, 1.6×, or 2× x-width (Figure 1F). Subjects were required to report the target letter and ignore the flankers.

The five conditions with different target–flanker spacing and one no-flanker condition were tested in a random order, each in a 20-trial block. The posttest followed the same spacing order as the pretest. The resulting accuracy-versus-spacing curve was then fitted with a psychometric function (cumulative Gaussian) to determine the spatial extent of crowding. Following Chung (2007), the spatial extent was defined as the letter separation (in multiples of x-width) yielding 50% recognition accuracy of the target letter, corrected for guessing. The reduction of crowding after training was quantified as the percentage change in the spatial extent of crowding. For instance, if the spatial extent was 1.5× x-width in the pretest and 1× in the posttest, then the reduction was $(1.5 − 1)/1.5 = 33\%$.

Training

Training consisted of six daily sessions from days 3 to 8, approximately 1 hr/day. For all three groups, the training task was to identify target letters, either flanked on both sides by letters or in isolation, at 10° in the lower visual field (Figure 2). One group was trained with crowded, localized letters (flanked-local group): Target letters always appeared in slot 0 for 150 ms, simultaneously flanked by two letters at a center-to-center spacing of 0.8× x-width. The subjects typed the center target letter after the stimuli disappeared. The second group was trained with crowded, distributed letters (flanked-distributed group): Target letters could appear for 150 ms in any slot from −6 to 6, and the center-to-center spacing between slots was standard spacing as in normal text (1.16× x-width, defined in the same way as in Figure 1B). The target letter was flanked by two letters at a separation of 0.8× x-width (as a result, the flankers did not fall exactly in the letter slots). Again, the subjects typed the center target letter. The third group was trained with isolated, distributed letters (isolated-distributed group): Target letters could appear in any slot from −6 to 6 (standard center-to-center spacing between slots) with no flanking letters but with reduced contrast and shorter exposure duration to increase task difficulty. The subjects typed the letter that appeared. Contrast level and the length of exposure duration were chosen for each individual

		Flanked-local	Flanked-distributed	Isolated-distributed
Spatial extent of crowding (multiples of *x*-width)		−25.4% ± 3.9%*	−11% ± 5.3%*	−7.2% ± 3.5%
Size of trigram visual span (bits)		+4.9 ± 0.6*	+4.7 ± 0.7*	+3.0 ± 0.8*
Reading performance	MRS (wpm)	+9.2% ± 5.0%	+30.3% ± 7.8%*	+29.9% ± 7.4%*
	CPS (°)	−13.5% ± 3.5%*	−10.2% ± 5.9%*	−13.4% ± 5.3%*

Table 2. Average changes from pre- to posttest (mean ± *SEM*). *Notes*: Significant changes are marked with an asterisk (*) as indicated by ANOVA or interaction analyses (see text). MRS = maximum reading speed. CPS = critical print size.

prior to training to achieve roughly 50%–70% correct during practice so that the task was below ceiling and well above chance.

Each daily training session had 10 blocks. For the flanked-local group, each block had 100 trials. For the flanked-distributed and isolated-distributed groups, each block had 104 trials, in which the target letter fell into each one of the 13 slots eight times.

Data analysis

When comparing training effects between groups, if not otherwise specified, we performed 3 × 2 mixed-design ANOVAs with group (flanked-local, flanked-distributed, isolated-distributed) as the between-subjects factor and session type (pre-/posttest) as the within-subject factor. If a significant interaction was found, we further analyzed the interaction using *R* with the package *phia* (post hoc interaction analysis; Martínez, 2015). The reported *p* values were adjusted for multiple comparisons within each analysis.

For MRS, CPS, and spatial extent of crowding, ANOVA analyses were performed using log-transformed data because the original curve fitting (two-limb function for reading and psychometric function for crowding) was performed on a log scale. When reporting pre- and posttest data, we have transformed the group-averaged log values back into their original units for easier understanding. The amount of percentage change from pre- to posttest (such as in Table 2) was computed individually using the original unit (say, wpm) and then averaged.

Results

Our main results are summarized in Figure 3 and Table 2. Panels A through D in Figure 3 show averaged group data with curves fitted to the average values. These curves are for demonstrative purposes only; data analyses were based on fitted values for individual subjects instead of the group-level curves (see Appendix 2 for individual curves). Panels E through G show

changes in key parameters from pre- to posttest for the three training groups, and the error bars indicate ±1 *SEM* for within-subject ANOVAs (Cousineau, 2005). The error bars can be used as a visual guide for within-group comparisons (between the pre- and posttests) but are not appropriate for between-group comparisons.

In the following sections, we first report the improvement on the trained task; then discuss how training affected the spatial extent of crowding, visual span, and reading separately; and last, put them together in a common framework in order to understand the nature of their associations.

Training progress

From Figure 3A, all three groups had improved performance on the trained task. The average accuracy of letter recognition improved from 0.48 to 0.65 from the first to the last training block and from 0.54 to 0.65 from the first to the last training day. The slopes of the linear fits (accuracy against training block) for all three groups were significantly larger than zero ($p < 0.001$), indicating significant improvement as training progressed.

Reduction of the spatial extent of crowding

Panels B and E of Figure 3 and the first row in Table 2 summarize the results for crowding measurements. The crowding curve (accuracy against letter spacing) for the flanked-local group shifted leftward after training, indicating improved recognition accuracy at smaller letter separations. The flanked-distributed group also showed some improvement although to a lesser extent. No improvement was apparent from the crowding curves of the isolated-distributed group.

We then derived the spatial extent of crowding (letter separation yielding 50% recognition accuracy, corrected for guessing) for each subject and performed an ANOVA to examine the effect of training (see interaction plot). We found a significant main effect of session type, $F(1, 24) = 33.25$, $p < 0.001$, as well as a significant interaction, $F(2, 24) = 4.98$, $p = 0.02$.
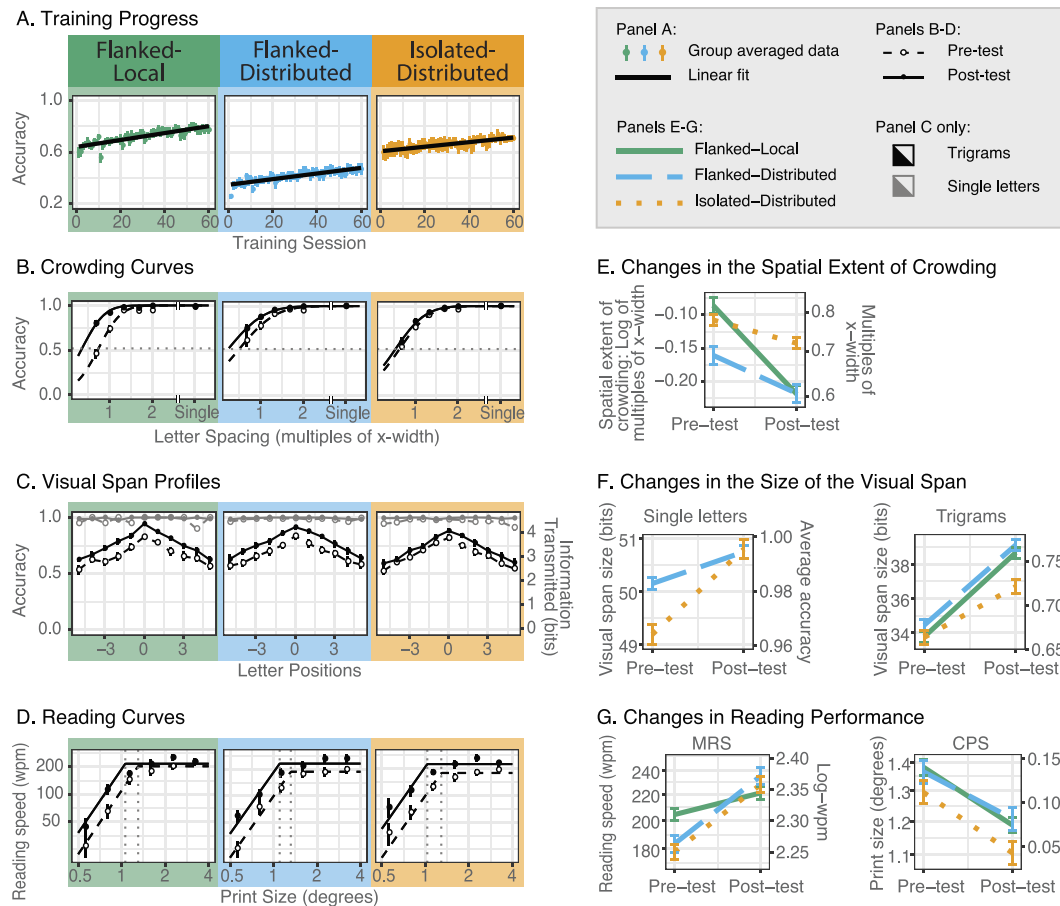
Figure 3. Summary of results. (A) Training progress (letter-recognition accuracy as a function of training session). (B–D) Pre–post comparison of three tests. Open circles and dashed lines: pretest; filled circles and solid lines: posttest. (B) Crowding curves (letter-recognition accuracy as a function of letter spacing). Solid and dashed black lines are cumulative Gaussian functions fitted to the group-averaged data. Dotted gray lines indicate 52% recognition accuracy (i.e., 50% accuracy after correction for guessing), which is the criterion used to determine the spatial extent of crowding. (C) Visual span profiles (letter-recognition accuracy as a function of letter positions) for single letters (gray) and trigrams (black). (D) Reading curves (reading speed as a function of print size on a log–log scale). Solid and dashed black lines are two-limb functions fitted to the group-averaged data. (E–G) Interaction plots showing changes in key parameters from pre- to posttest for the three training groups. Green solid lines: flanked-local group; blue dashed lines: flanked-distributed group; orange dotted lines: isolated-distributed group. (E) Changes in the spatial extent of crowding. (F) Changes in the size of the visual span for single letters and trigrams. (G) Changes in MRS and CPS. In panels A–D, error bars indicate ± 1 SEM. In panels E–G, error bars indicate ± 1 SEM for within-subject ANOVAs (Cousineau, 2005).

Analysis of interaction showed that only the two flanked groups had statistically significant reduction of crowding after training (flanked-local: spatial extent of crowding decreased from 0.82 to 0.60 times x-width, mean reduction 25.4%, $p < 0.001$, adjusted for multiple comparisons; flanked-distributed: from 0.69 to 0.61 times x-width, average reduction 11%, adjusted $p = 0.03$) but not the isolated-distributed group (from 0.78 to 0.72 times x-width, average reduction 7.2%, adjusted $p = 0.13$). The reduction for the flanked-local group was significantly larger than that for the flanked-distributed group (adjusted $p = 0.04$) and the isolated-distributed group (adjusted $p = 0.008$). No difference was found between the two distributed groups (adjusted $p = 0.50$).

One concern is that the pretest performance of the flanked-distributed group was poorer than the other two groups, which may leave more room for improvement. We therefore performed a linear regression between the reduction of crowding (difference in spatial extent, log unit) and the pretest crowding level. The regression line had a negative slope of -0.47 that was significantly different from zero ($p = 0.009$); i.e., larger reduction was associated with poorer pretest level. This indicates that in the previous analysis we had underestimated the reduction of crowding in the flanked-distributed group because it had a better starting level than the other two groups. A new between-group ANOVA was performed on the reduction of crowding, using the pretest level as a covariate. After accounting

for the pretest level, there was still a main effect of group, $F(2, 23) = 4.62$, $p = 0.02$. A post hoc analysis showed that crowding was only reduced in the two flanked groups (adjusted $p < 0.01$) but not in the isolated group ($p = 0.14$). When comparing between groups, a slightly different pattern emerged: When comparing the two flanked groups, although the face value of the reduction of crowding was larger for the flanked-local group, the difference did not reach significance after pretest level was taken into account (adjusted $p = 0.28$). The reduction for the flanked-local group was significantly larger than for the isolated-distributed group (adjusted $p = 0.007$), and the two distributed groups did not differ (adjusted $p = 0.28$).

Taken together, our analyses suggest that the flanked-local training was the most effective in reducing the spatial extent of crowding, followed by the flanked-distributed training, but the isolated-distributed training was not effective. The isolated-distributed group can be viewed as a sham training group, and the improvement sets an upper bound on the test–retest effect.

Enlargement of the visual span

Panels C and F of Figure 3 and the second row of Table 2 summarize the results for visual span measurement. For single-letter visual span profiles (gray lines and symbols), only a very small change was observed. For trigrams (black lines and symbols), all groups exhibited notable training-related enlargements, and the improvement appeared to be smaller for the isolated-distributed group compared to the other two groups.

We then performed ANOVAs on visual span sizes for single letters and trigrams, respectively. For single letters, we only used the data from the flanked-distributed and isolated-distributed groups because in the flanked-local group only one subject had both the pre- and posttest data in this condition. We found a significant main effect of group, $F(1, 16) = 6.88$, $p = 0.019$; a marginal significant main effect of session type, $F(1, 16) = 3.75$, $p = 0.07$; and a significant interaction between the two factors, $F(1, 16) = 5.38$, $p = 0.034$. An interaction analysis revealed that only the isolated-distributed group had statistically significant improvement after training (from 49.2 to 50.8 bits or averaged accuracy from 96.4% to 99.5%; adjusted $p < 0.001$), and this improvement was significantly larger ($p = 0.02$) than that for the flanked-distributed group (from 50.1 to 50.7 bits or averaged accuracy from 98.3% to 99.4%). Nevertheless, the absolute value of the improvement was very small due to the ceiling effect and thus may not reflect large changes in visual functions.

An ANOVA on the trigram data revealed a main effect of session type, $F(1, 24) = 48.7$, $p < 0.001$, indicating enlarged visual spans after training. No main effect of group or interaction was found. Although the groups were not significantly different, the absolute enlargement for the isolated-distributed group (from 33.7 to 35.7, +3 bits, or averaged accuracy from 66.4% to 72.1%) was smaller than that for the other two groups (flanked-local group: from 33.8 to 38.7, +4.9 bits, or averaged accuracy from 66.4% to 76.0%; flanked-distributed group: from 34.4 to 39.1, +4.7 bits, or averaged accuracy from 67.8% to 76.8%). This pattern is also apparent from the interaction plot in Figure 3F.

This enlargement evaluates changes accumulated over 11 letter slots (from −5 to 5) on the visual span profile. Past studies using similar trigram training have found enlargement of 5.4 bits accumulated over nine letter slots (He et al., 2013) or 6.1 bits over 11 slots (Chung, Legge, & Cheung, 2004) whereas their corresponding no-training control groups had 0.5 bits (He et al., 2013) and about 1.2 bits change (estimated from figure 9B in Chung et al., 2004). Our training-related enlargement is smaller compared to previous training paradigms but is larger than no-training controls. We return to this difference in the Discussion.

Improvement in MRS

Reading performance is summarized in Panels D and G of Figure 3 and the bottom two rows of Table 2. We focus on the change of MRS in the following discussion and report the results on CPS in the Appendix.

As shown by the average reading curves in Figure 3D, only the two distributed groups had larger MRS after training. For the flanked-local group, MRS remained almost unchanged. An analysis of individual fitted parameters confirmed the group-level pattern as shown in the interaction plot. A 3×2 ANOVA on log MRS revealed no main effect but a marginally significant interaction between group and session type, $F(2, 24) = 3.15$, $p = 0.06$. Analysis of the interaction showed that only the two distributed groups showed significant improvement after training (flanked-distributed: from 183 to 235 wpm, average improvement 30.3%; isolated-distributed: from 178 to 228 wpm, average improvement 29.9%; both of their adjusted ps < 0.001). MRS for the flanked-local group changed from 204 to 221 wpm, which is not statistically significant (a change of 9.2%; adjusted $p = 0.15$). Further between-group comparisons showed that the flanked-local group had marginally less improvement in MRS than the other two distributed groups (both of the adjusted ps $= 0.087$), but the two distributed groups were not significantly different (adjusted $p = 0.99$). Our

design did not include a no-training control group, but control data were available from a similar training study (Chung et al., 2004). In their study, trigram-recognition training improved MRS by 41%, but without training, MRS increased by about 6% (estimated from figure 9C in Chung et al., 2004). The improvement for the flanked-local group was similar to their no-training control group.

However, we again have the concern that the difference in the pretest level may produce misleading group differences. We therefore performed a linear regression of the pre–post difference in log MRS against the pretest log MRS level. This time, the improvement did not depend on the pretest level; i.e., the slope of the regression line is not significantly different from zero ($p = 0.44$). This suggests that higher pretest performance level cannot account for the lack of improvement in the flanked-local group.

Discussion

Summary of results

We used three different training procedures in which the training tasks differed in a systematic way, in order to test the influence of (a) spatial distribution of training stimuli and (b) the presence or absence of flanking letters. As observed in Table 2 and summarized from the above analyses, we found that:

- The spatial extent of crowding only decreased in the two flanked groups but not in the isolated group, suggesting that flanked training stimuli are necessary to reduce the spatial extent of crowding. Moreover, the reduction of crowding was larger in the flanked-local group than the flanked-distributed group. That is, when all crowded training stimuli are located at one single location, the training effect for that location is larger than when crowded stimuli are distributed across spatial locations. This reflects some degree of location specificity in the improvement.
- The size of the trigram visual span improved for all three groups, and the enlargement was smaller for the isolated group. This suggests that the visual span enlarges after practicing with letter recognition, no matter flanked or isolated, distributed or localized, although the use of flanked training stimuli may result in a greater improvement.
- MRS improved in the two distributed training groups while the local training group showed small changes similar to a no-training control. This suggests that the transfer of training to reading speed is limited when training stimuli have fixed

spatial locations. For the conditions we have tested, training with distributed targets is both necessary and sufficient for such transfer to occur. Conversely, the presence of flanking letters is neither a necessary nor sufficient condition for the transfer.

Reduction of the spatial extent of crowding

Our results suggest that flanked training stimuli are necessary to reduce the spatial extent of crowding. This was expected because there was no crowding in isolated training. However, others have found that training with isolated letters can sometimes reduce the impact of crowding. In adults with amblyopia, training with isolated, near-acuity, reduced-contrast letters reduced the spatial extent of foveal crowding (Chung, Li, & Levi, 2012). Amblyopic vision shares similarities with normal peripheral vision, including properties of crowding (for example, see the review in Levi et al., 2007). If isolated letters can reduce crowding for amblyopic vision, it is plausible that similar benefits could occur in normal peripheral vision.

Our training with isolated letters did not result in a reduced spatial extent of crowding, possibly because our training was distributed across various letter positions, whereas Chung et al. (2012) trained subjects at one fixed location. As suggested by the comparison between our flanked-local and flanked-distributed groups, and consistent with previous training studies (Yashar, Chen, & Carrasco, 2015; Zhu, Fan, & Fang, 2016), the reduction in the spatial extent of crowding had some degree of location specificity. Taken together, learning to uncrowd a localized target benefited more from localized and crowded training than distributed or isolated training.

Enlargement of the visual span

We found that the visual span enlarged with letter-recognition training regardless of the distribution of the stimuli or the presence of flankers. One surprising result was that when training with isolated letters, the spatial extent of crowding did not change whereas the visual span enlarged. Given the association between crowding and visual span outlined in the Introduction, this result was unexpected.

But our results do not sever the proposed link between crowding and visual span. Undoubtedly, the size of the visual span is mainly limited by the interfering effect on letter recognition from nearby letters. However, the spatial extent of crowding at one single letter position may not be enough to characterize such an effect. First, letter positions farther away from

the midline are more relevant for determining the boundary of the visual span. Pelli et al. (2007) found that at the boundary of the visual span, letter spacing is equal to the critical spacing of crowding. Second, in addition to the extent of crowding, its amplitude also needs to be considered. As discussed by Pelli and Tillman (2008, supplementary discussion), crowding can be weakened in its magnitude without changing its extent. Our isolated training may have reduced the extent of crowding at locations other than the midline or reduced the magnitude rather than the extent of crowding, resulting in the decoupling of the spatial extent of crowding and the size of the visual span. In future research, the choice of measurement should distinguish between the spatial extent and magnitude of crowding.

If training with both isolated and crowded letters can enlarge the visual span, what might be the underlying mechanism? One possibility is a better template-matching process for letter recognition. Training may modify the perceptual templates used by human observers to be more similar to the templates used by a theoretical ideal observer (Gold, Sekuler, & Bennett, 2004) as well as adjust the spatial extent of the perceptual window used for sampling (Sun, Chung, & Tjan, 2010). If our training resulted in a more appropriate sampling window for letter recognition and/or improved templates of letters or important features for letter recognition, such as line terminations (Fiset et al., 2008), the visual span would enlarge. However, only a better sampling window could lead to a reduced spatial extent of crowding. In our isolated-distributed group, it seems that training only improved perceptual templates without adjusting the size of the sampling window.

The enlargement of the visual span was similar for the flanked-local group (4.9 bits) and the flanked-distributed group (4.7 bits) despite the fact that the local training only had limited retinotopic overlap (roughly three letter slots) with the visual span measurement (13 letter slots). The lack of location specificity in the training effect was consistent with previous findings in which training in the upper (or lower) hemifield enlarged visual spans in the untrained hemifield (for example, Chung et al., 2004; He et al., 2013; Lee, Kwon, Legge, & Gefroh, 2010; Yu, Legge, Park, Gage, & Chung, 2010). This suggests a non-retinotopic-specific mechanism underlying the enlargement of the visual span, such as the better template-matching process discussed before.

One potential factor influencing the enlargement of the visual span is the partial report method we used during training. In our training paradigm, only one (middle) letter needed to be reported. As a result, attention to the flanking letters was likely diminished. Consistently, the enlargement in our study was smaller than in some similar studies using full-report trigram training (e.g., Chung et al., 2004; He et al., 2013).

Improvement in MRS

Our major finding is that for the training benefit to transfer from letter recognition to increased reading speed, spatially distributed training stimuli appear to be necessary. This result is consistent with previous findings in which MRS improved with distributed training (41% improvement with trigram measurement training, Chung et al., 2004) but not with localized training (7.2% change with "uncrowd" training similar to our flanked-local group, Chung, 2007). With our distributed training, 1-bit enlargement of visual span on average corresponded to 0.028 log wpm improvement in reading speed (or 6.7% improvement), close to the previously reported value of 0.03 log wpm/bit averaged from various studies (Legge et al., 2007). In the following paragraphs, we discuss how spatially distributed training differs from localized training in its influence on reading. We focus on two aspects: retinotopic overlap and the deployment of attention during the task.

First, we consider the retinotopic overlap of training stimuli (letters) and reading test stimuli (words). Localized training stimuli were constrained to a very small horizontal span, occupying on average 5.7° in width. In the reading test, this span was equivalent to about one to three letter slots for the largest three print sizes for which MRS was reached (Figure 3, bottom left). Therefore, there was only limited retinotopic overlap between localized training and the words in larger print sizes, possibly making localized training less effective in improving MRS compared to distributed training. However, insufficient retinotopic exposure cannot fully explain the lack of improvement because the training effect in reading speed has been shown to transfer between upper and lower hemifields (e.g., Chung et al., 2004; He et al., 2013; Lee et al., 2010; Yu et al., 2010). This suggests a higher level, nonretinotopic mechanism underlying the lack of improved MRS in the present study.

A more probable explanation for the lack of benefit of localized training is the deployment of visual attention. Localized training may facilitate the deployment of attention to one specific, small area in the visual field. For instance, in a visual search task, if the target more frequently appears in a certain spatial area, attention is implicitly biased toward that area (Geng & Behrmann, 2002). It only takes dozens of trials to develop such a bias, but once learned, the bias is not easily unlearned and can last for at least a week (Jiang, Swallow, Rosenbaum, & Herzig, 2013). Similarly, during localized training for our flanked-local group, subjects may have acquired a sustained bias of spatial

attention toward the letter position right below fixation. This attentional bias is beneficial for reducing the spatial extent of crowding at a specific location, but during reading, it may limit the spread of attention to the letters of longer words, negating potential benefit from enlarged visual span.

In addition, as discussed in the previous section, the partial report method in the flanked training may have introduced an additional attentional bias with which the attention to flankers was discouraged. If this attentional bias transferred to reading, it would hamper the parallel recognition of letters necessary for rapid reading. In support, our training using partial report achieved about 30% improvement in MRS whereas similar trigram training studies using full-report resulted in larger improvement (40%–66%; Chung et al., 2004; He et al., 2013; Lee et al., 2010; Yu et al., 2010).

In normal eye movement–based reading, the allocation of attention between words is often discussed in models of eye movement control, for example, the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998) and the SWIFT model (Engbert, Longtin, & Kliegl, 2002). In such models, the deployment of attention to parafoveal words facilitates subsequent lexical processing when the words later become fixated. Impaired ability in distributing visual spatial attention, such as sluggish rightward shift of attention, has been associated with reading deficits in people with dyslexia (Schneps et al., 2013). In RSVP reading, although the eye-movement models do not apply, the allocation of attention within a word is still important for reading. For example, if attention is distributed to one letter within a word rather than across the whole word, semantic processing of that word will be impaired (for a review, see Besner et al., 2016). Together with our results, these pieces of evidence indicate an important role for distributed attention in reading.

Linking crowding, visual span, and reading

Now we return to the link between crowding, visual span, and reading.

The link between crowding and visual span seems to break when training with distributed, isolated letters with which the spatial extent of crowding remained unchanged but visual span enlarged. As we previously discussed, in the case of the isolated-distributed group, training reduced the impact of crowding on the size of the visual span without inducing measurable changes in its spatial extent. In the context of the measurement of visual-span profiles with trigram stimuli, crowding refers to the effect of flankers at a fixed spacing on the accuracy of letter identification. Reduced strength of crowding,

rather than reduced spatial extent of crowding, may have accounted for the enlargement of visual span.

The link between visual span and reading seems to break when training with localized, crowded letters, with which the visual span enlarged but MRS remained unchanged. We proposed that a disadvantageous attentional bias toward a small area limits the improvement in reading. During rapid sequential presentation of words, such a bias will likely direct attention to a very narrow region within the word, resulting in recognition of only a limited number of letters under time pressure. In contrast, during the measurement of visual span, no matter where the trigram appears, the abrupt and brief appearance of the trigram is a strong exogenous cue. This cue will automatically disengage attention from the previously prioritized location and reorient it to the location of the trigram. The low task demand (only three letters need to be reported rather than a sequence of words) and the unlimited response time further alleviate potential impact of the attentional bias. Therefore, the attentional bias slows down reading but only has minimal impact on visual span. Reading is still limited by the size of the visual span but also depends on the spatial distribution of attention.

We conclude that the visual span represents a sensory bottleneck on reading, but there may also be an attentional bottleneck. Crowding affects reading by limiting the size of the visual span. Reducing the impact of crowding can enlarge the visual span and can potentially facilitate reading, but for some types of training, the benefits for reading may be offset by an attentional bias. Our results thus clarify the association between crowding, visual span, and reading.

Commercial relationships: Precision Vision (GEL) (R).
Corresponding author: Yingchen He.
Email: hexxx340@umn.edu.
Address: Department of Psychology, University of Minnesota, Twin Cities, MN, USA.

References

Besner, D., Risko, E. F., Stolz, J. A., White, D., Reynolds, M., O'Malley, S., & Robidoux, S. (2016).

Varieties of attention. *Current Directions in Psychological Science*, *25*(3), 162–168, doi:10.1177/0963721416639351.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. Retrieved from http://bbs.bioguider.com/images/upfile/2006-4/200641014348.pdf

Callens, M., Whitney, C., Tops, W., & Brysbaert, M. (2013). No deficiency in left-to-right processing of words in dyslexia but evidence for enhanced visual crowding. *Quarterly Journal of Experimental Psychology*, *66*(9), 1–33, doi:10.1080/17470218.2013.766898.

Cheong, A. M. Y., Legge, G. E., Lawrence, M. G., Cheung, S.-H., & Ruff, M. A. (2007). Relationship between slow visual processing and reading speed in people with macular degeneration. *Vision Research*, *47*(23), 2943–2955, doi:10.1016/j.visres.2007.07.010.

Cheung, S.-H. H., Kallie, C. S., Legge, G. E., & Cheong, A. M. Y. (2008). Nonlinear mixed-effects modeling of MNREAD data. *Investigative Ophthalmology & Visual Science*, *49*(2), 828–835. [PubMed] [Article]

Chung, S. T. L. (2002). The effect of letter spacing on reading speed in central and peripheral vision. *Optometry and Vision Science*, *43*(4), 1270–1276, doi:10.1097/00006324-200012001-00034.

Chung, S. T. L. (2007). Learning to identify crowded letters: Does it improve reading speed? *Vision Research*, *47*(25), 3150–3159, doi:10.1016/j.visres.2007.08.017.

Chung, S. T. L., Legge, G. E., & Cheung, S.-H. (2004). Letter-recognition and reading speed in peripheral vision benefit from perceptual learning. *Vision Research*, *44*(7), 695–709, doi:10.1016/j.visres.2003.09.028.

Chung, S. T. L., 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.

Chung, S. T. L., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research*, *38*(19), 2949–2962. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9797990

Chung, S. T. L., & Truong, S. R. (2013). Learning to identify crowded letters: Does the learning depend on the frequency of training? *Vision Research*, *77*, 41–50, doi:10.1016/j.visres.2012.11.009.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology*, *1*(1), 42–45.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262, doi:10.1016/j.tics.2011.04.003.

Doron, A., Manassi, M., Herzog, M. H., & Ahissar, M. (2015). Intact crowding and temporal masking in dyslexia. *Journal of Vision*, *15*(14):13, 1–17, doi:10.1167/15.14.13. [PubMed] [Article]

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*(5), 621–636, doi:10.1016/S0042-6989(01)00301-7.

Fiset, D., Blais, C., Éthier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for uppercase and lowercase letter identification. *Psychological Science*, *19*(11), 1161–1168.

Forster, K. I. (1970). Visual perception of rapidly presented word sequences of varying complexity. *Perception & Psychophysics*, *8*(4), 215–221, doi:10.3758/BF03210208.

Geng, J. J., & Behrmann, M. (2002). Probability cuing of target location facilitates visual search implicitly in normal participants and patients with hemispatial neglect. *Psychological Science*, *13*(6), 520–525, doi:10.1111/1467-9280.00491.

Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, *28*(2), 167–207, doi:10.1016/j.cogsci.2003.10.005.

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]

Jiang, Y. V., Swallow, K. M., Rosenbaum, G. M., & Herzig, C. (2013). Rapid acquisition but slow extinction of an attentional bias in space. Journal of Experimental Psychology: Human Perception and Performance, *39*(1), 87–99, no pagination specified, doi:10.1037/a0027611.

Lee, H.-W., Kwon, M., Legge, G. E., & Gefroh, J. J. (2010). Training improves reading speed in peripheral vision: Is it due to attention? *Journal of Vision*, *10*(6):18, 1–15, doi:10.1167/10.6.18. [PubMed] [Article]

Legge, G. E., Cheung, S.-H., Yu, D., Chung, S. T. L., Lee, H., & Owens, D. P. (2007). The case for the visual span as a sensory bottleneck in reading.

Journal of Vision, 7(2):9, 1–15, doi:10.1167/7.2.9. [PubMed] [Article]

Levi, D. M., Song, S., & Pelli, D. G. (2007). Amblyopic reading is crowded. *Journal of Vision*, 7(2):21, 1–17, doi:10.1167/7.2.21. [PubMed] [Article]

Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *Journal of Vision*, 9(4):14, 1–18, doi:10.1167/9.4.14. [PubMed] [Article]

Martínez, H. D. R. (2015). Analysing interactions of fitted models. The Comprehensive R Archive Network. Retrieved from http://cran.wustl.edu/web/packages/phia/vignettes/phia.pdf

Moll, K., & Jones, M. (2013). Naming fluency in dyslexic and nondyslexic readers: Differential effects of visual crowding in foveal, parafoveal, and peripheral vision. *Quarterly Journal of Experimental Psychology*, 66(11), 2085–2091, doi:10.1080/17470218.2013.840852.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442, doi:10.1163/156856897X00366.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 11(10), 1129–1135, doi:10.1038/nn.2187.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision*, 7(2):20, 1–36, doi:10.1167/7.2.20. [PubMed] [Article]

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105(1), 125–157, doi:10.1037/0033-295X.105.1.125.

Rubin, G. S., & Turano, K. (1992). Reading without saccadic eye movements. *Vision Research*, 32(5), 895–902. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16005930

Schneps, M. H., Thomson, J. M., Sonnert, G., Pomplun, M., Chen, C., & Heffner-Wong, A. (2013). Shorter lines facilitate reading in those who struggle. *PloS One*, 8(8), e71161, doi:10.1371/journal.pone.0071161.

Sun, G. J., Chung, S. T. L., & 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]

Yashar, A., Chen, J., & Carrasco, M. (2015). Rapid and long-lasting reduction of crowding through training. *Journal of Vision*, 15(10):15, 1–15, doi:10.1167/15.10.15. [PubMed] [Article]

Yu, D., Cheung, S.-H., Legge, G. E., & Chung, S. T. L. (2007). Effect of letter spacing on visual span and reading speed. *Journal of Vision*, 7(2):2, 1–10, doi:10.1167/7.2.2. [PubMed] [Article]

Yu, D., Legge, G. E., Park, H., Gage, E., & Chung, S. T. L. (2010). Development of a training protocol to improve reading performance in peripheral vision. *Vision Research*, 50(1), 36–45, doi:10.1016/j.visres.2009.10.005.

Zhu, Z., Fan, Z., & Fang, F. (2016). Two-stage perceptual learning to break visual crowding. *Journal of Vision*, 16(6):16, 1–12, doi:10.1167/16.6.16. [PubMed] [Article]

Appendix 1: CPS

From the averaged reading curves in Figure 3D, all three groups had reduced CPS (vertical dashed lines) after training. A 3×2 ANOVA was performed for CPS after log transformation with group (flanked-local, flanked-distributed, isolated-distributed) as the between-subjects factor and session type (pre-/posttest) as the within-subject factor. We found a significant main effect of session type, $F(1, 24) = 7.67$, $p = 0.01$: overall, CPS decreased from 1.35° to 1.17° after training (average reduction 12.3%). No main effect of group or any interaction effect was found.

Our results indicate that CPS decreased no matter whether the training was localized or distributed. If the attentional bias limits the improvement of MRS, why does it not limit the reduction of CPS? Note that unlike MRS, the reduction of CPS is more closely related to the improvement in reading speed for smaller-sized text. When text size is smaller, even a localized training will have large retinotopic overlap with most words. For example, in our study, the mean training size for the localized training group (1.4 times the subjects' average CPS) was 1.85° in x-height whereas the two smallest sizes in the reading test were 0.56° and 0.79° in x-height. On average, training stimuli (three letters spaced 0.8 times x-width) spanned 5.6° horizontally. The average word length of our reading test material was four letters, and a four-character word would span 3° when x-height is 0.56° or 4.3° when x-height is 0.79°. Therefore, our training has enough coverage to stimulate the letters in most of the words for the two smallest print sizes in our reading task. In this case, even if a spatial bias was introduced during training, the preferred location largely overlaps with the words, and thus, the bias does not have a detrimental effect.

One issue may cast doubt on the above discussion. We found seemingly inconsistent results with a previous study: Using similar distributed training, we found reduced CPS after training whereas Chung et al. (2004) did not. But the results are not conflicting if we focus on the improvement in reading smaller-sized text instead of on CPS per se. It seems that in Chung et al. (2004), reading speed (in log wpm) improved uniformly across all print sizes, resulting in increased MRS but no change in CPS. As can be inferred from the reading curve in Figure 1E, if reading speed improved uniformly across all print sizes, CPS would remain unchanged despite the improvement in reading small-sized text. This is the case in Chung et al. (2004) in which improvement in reading smaller-sized text did not lead to a reduced CPS. Our results are consistent with Chung et al. (2004) and Chung (2007) in that both distributed and localized training can improve reading speed for small-sized text.

In summary, CPS decreased in all three training groups. This indicates that neither crowded training stimuli nor spatially distributed training was necessary for this decrease.
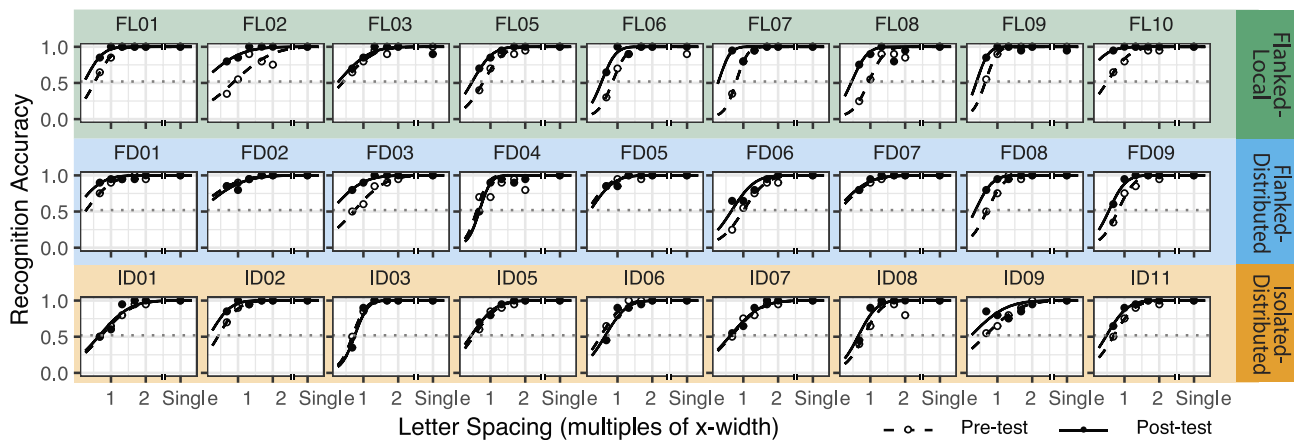
Appendix 2: Individual figures
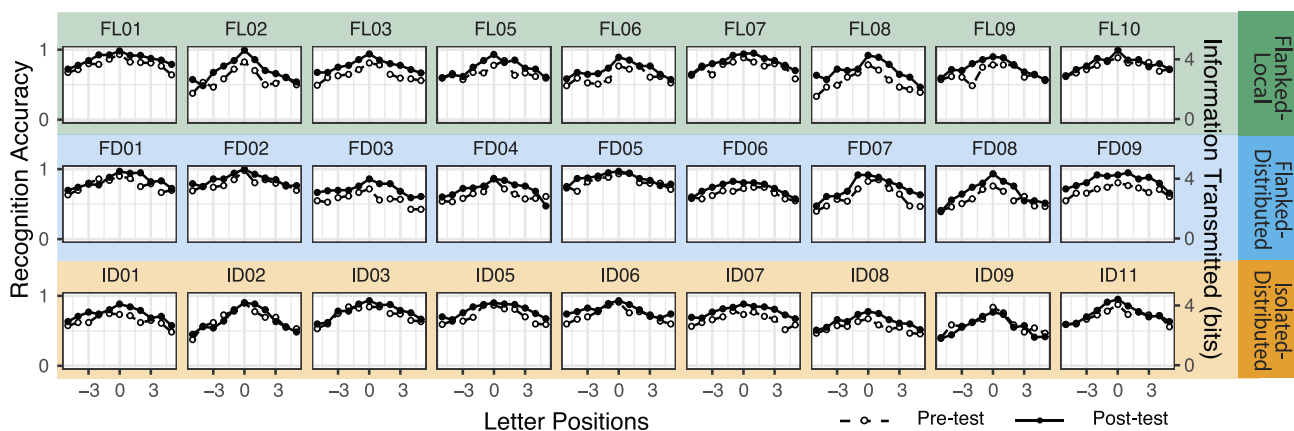

Figure A2-1. Individual crowding curves.
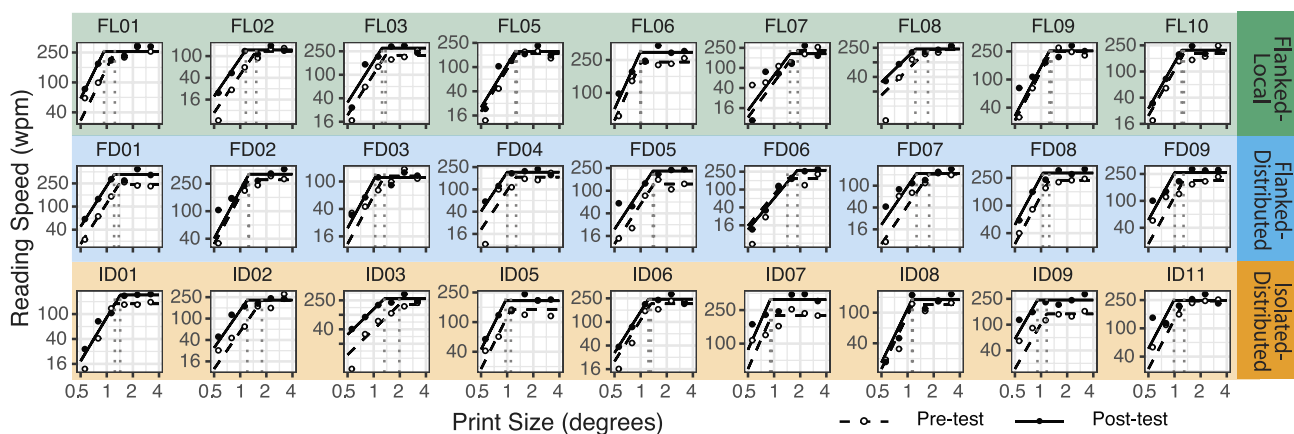

Figure A2-2. Individual trigram visual span profiles.

Figure A2-3. Individual reading curves.