Available online at www.sciencedirect.com

ScienceDirect

Vision Research 47 (2007) 3150–3159

Vision Research

www.elsevier.com/locate/visres

Learning to identify crowded letters: Does it improve reading speed?

Susana T.L. Chung *

College of Optometry, University of Houston, Houston, TX, USA

Received 10 June 2007; received in revised form 19 August 2007

Abstract

Crowding, the difficulty in identifying a letter embedded in other letters, has been suggested as an explanation for slow reading in peripheral vision. In this study, we asked whether crowding in peripheral vision can be reduced through training on identifying crowded letters, and if so, whether these changes will lead to improved peripheral reading speed. We measured the spatial extent of crowding, and reading speeds for a range of print sizes at 10° inferior visual field before and after training. Following training, averaged letter identification performance improved by 88% at the trained (the closest) letter separation. The improvement transferred to other untrained separations such that the spatial extent of crowding decreased by 38%. However, averaged maximum reading speed improved by a mere 7.2%. These findings demonstrated that crowding in peripheral vision could be reduced through training. Unfortunately, the reduction in the crowding effect did not lead to improved peripheral reading speed.

Keywords: Crowding; Perceptual learning; Training; Reading

1. Introduction

Reading is difficult and slow for many low vision patients, especially those whose central retina is damaged, and consequently suffer from central vision loss. The leading cause of central vision loss is age-related macular degeneration, which is also the leading cause of visual impairment in developed countries (Elliott et al., 1997; Leat & Rumney, 1990). A consequence of the central vision loss is the loss of acute vision, which affects visual tasks that involve fine details such as reading. Given that reading is the most common clinical complaint as well as the primary goal for patients with age-related macular degeneration seeking visual rehabilitation (Bullimore & Bailey, 1995; Elliott et al., 1997; Kleen & Levoy, 1981), the understanding of why reading is slower in the presence of central vision loss is of utmost importance to the visual rehabilitation of these patients.

Previous studies have shown that even when character size is not a limiting factor (Chung, Mansfield, & Legge, 1998; Latham & Whitaker, 1996), and when oculomotor demands for reading are minimized using the rapid serial visual presentation (RSVP) paradigm to present text, reading is still slower in peripheral than central vision (Chung et al., 1998; Latham & Whitaker, 1996; Rubin & Turano, 1994). Recently, Legge and his colleagues provided empirical evidence showing that the size of the visual span—the number of letters that can be recognized in a glance—correlates well with reading speed, implicating that the visual span may be a sensory bottleneck on reading speed (Legge et al., 2007). With respect to peripheral reading, the visual span hypothesis suggests that slow peripheral reading is due to a shrinkage in the size of the visual span (Legge, Mansfield, & Chung, 2001). However, what determines the size of the visual span? Legge et al. (2007) proposed that the size of the visual span is determined by three sensory mechanisms, *viz,* the decreased letter acuity in peripheral vision, crowding between adjacent letters and decreased accuracy of position signals in peripheral vision. Although the role of each of these factors cannot be

---

* Fax: +1 713 743 2053.
E-mail address: schung@optometry.uh.edu

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.08.017

ignored, in this paper, we shall focus on only one of them—crowding between adjacent letters.

The crowding phenomenon refers to "the increased difficulty in identifying targets which are closely adjacent to other targets" (Cline, Hofstetter, & Griffin, 1997, p. 521). With respect to letters or words, crowding refers to the difficulty in recognizing a letter flanked by other letters at small letter separations. It has been suggested as a viable factor contributing to slow reading in peripheral vision because even when letters are scaled in size, the spatial extent (Jacobs, 1979; Latham & Whitaker, 1996; Toet & Levi, 1992) and the magnitude (Jacobs, 1979; Loomis, 1978) of the effect are still greater in peripheral than central vision. However, does crowding really limit peripheral reading?

The classic study of Bouma (1970) established that the performance of identifying flanked letters improved with letter separation. Whittaker, Rohrkaste, and Higgins (1989) and Latham and Whitaker (1996) further showed that word recognition performance improved with increased letter separation. Based on this evidence, it was almost only logical to infer that reading speed would improve with letter separation. We tested this idea by measuring reading speed as a function of letter separation (Chung, 2002). Contrary to our expectation, we found that reading speed did not improve with larger-than-normal letter separation. This finding has since been replicated by at least two other studies (Berger, Martelli, Su, Aguayo, & Pelli, 2003; Yu, Cheung, Legge, & Chung, 2007).

The finding that increased letter separation does not lead to faster reading speed might seem surprising. Yu et al. (2007) applied the visual-span concept to explain their data. They found that reading speed dropped for an extra-wide letter spacing (wider than the standard letter spacing), and it was associated with a reduction in the size of the visual span. Thus, they explained the failure of an increased letter separation, which presumably reduces crowding among adjacent letters, to improve reading speed as a consequence of a reduction in the size of the visual span. Pelli and his colleagues further postulated that the size of the visual span is limited by crowding, and further inferred that reading speed is also limited by crowding (Levi, Song, & Pelli, 2007; Pelli et al., 2007).

Undoubtedly, testing if reading speed improves with large letter separations is only one way to evaluate if crowding indeed limits reading speed in peripheral vision. In this study, we used another approach to examine the relationship between crowding and peripheral reading. If crowding indeed limits peripheral reading, then it is important to ask whether we can reduce crowding in peripheral vision, and if so, whether there is a corresponding improvement in peripheral reading speed. These were the primary questions we addressed in this study. Specifically, we examined (1) whether or not we could reduce crowding in peripheral vision by improving the performance for identifying crowded letters through repeated training on a letter recognition task; and (2) whether these improvements in letter identification were accompanied by comparable improvements in reading speed.

To our knowledge, there exists very little evidence showing that practice leads to an improved performance for identifying crowded letters. The only related study examined the effect of practice on contour interaction (Manny, Fern, Loshin, & Martinez, 1988). Contour interaction refers to the effect of proximal contours such as bars or edges on the resolution of a single target (Flom, Weymouth, & Kahneman, 1963) and is different from crowding (Flom, 1991). In the study of Manny et al. (1988), observers' performance for identifying the gap of a square-C pattern in the presence of flanking bars improved after 1200 trials of practice with feedback at the fovea. However, the authors showed that what the observers learned was to make discriminations based on the local luminance cues provided by the gap of the C patterns. These cues might be useful for their contour interaction task, but they are unlikely to be useful for identifying crowded letters. Therefore, it remains to be proven that performance for identifying crowded letters could be improved with practice, and specifically, in peripheral vision.

Previously, we have shown that letter identification for sequences of three letters, *trigrams*, improved with repeated training, and that the improvement was accompanied by a 41% improvement in reading speed in peripheral vision (Chung, Legge, & Cheung, 2004). In that study, the training task was to identify trigrams at a range of letter positions. An auxiliary purpose of the present study was to test if we could use a simpler training task to achieve the same amount of improvement. A simpler training task would be more appealing and practical because our ultimate goal is to develop a training protocol to improve visual function in patients with central vision loss. In this study, we chose to train observers intensively only at one letter position, instead of multiple letter positions as in our previous study (Chung et al., 2004).

To anticipate our major findings, we found that performance for identifying crowded letters in peripheral vision improves with training. This improvement transfers to other untrained letter separations, such that the spatial extent of crowding decreases following training. However, the improvement in letter identification performance is not accompanied by a sizeable improvement in reading speed, suggesting that crowding is unlikely to be the primary factor that limits reading speed in peripheral vision.

2. Methods

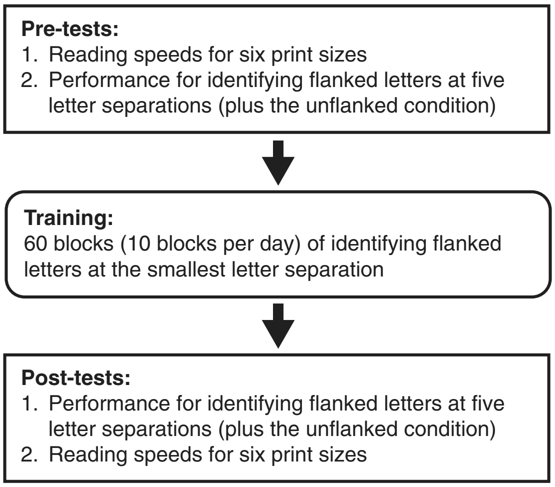
Eight observers with normal vision, aged 20–31 (mean = 23.1), participated in this study. All were native English speakers and had not participated in any psychophysical perceptual learning studies before. Written informed consent was obtained from each observer after the procedures of the experiment were explained and prior to the commencement of data collection. All observers participated in all three phases of this study—pre-test, training and post-test, that involved measurements of reading speed and letter identification performance at 10° in the inferior visual field.

The basic experimental design and training schedule are represented schematically in Fig. 1. The pre-test was conducted in two sessions, scheduled on two days. The first session was devoted to the measurement of reading speed as a function of print size and lasted approximately 1.5 h. From these data, we determined the appropriate letter size (see below for details) used for the second pre-test session during which we measured the accuracy of identifying flanked letters for a range of letter separations.

Because the second pre-test session was rather short (approximately 15 min), all observers proceeded to the first training session immediately after the second pre-test session. Each training session, lasting approximately 1 h, consisted of 10 blocks of trials (100 trials per block) of identifying flanked letters at the closest (0.8×, see Section 2.2 for definition) letter separation. There were six training sessions altogether (a total of 6000 trials), scheduled on six different days. The post-test immediately followed the last training session. It was identical to the pre-test except that the measurement of letter identification performance preceded the reading speed measurement, so that we could measure the performance for identifying flanked letters for a range of letter separations immediately before and after training. All observers completed all phases (pre-test, training and post-test) of the experiment within a 10-day period.

2.1. Reading speed measurements

The first pre-test and the second post-test sessions were devoted to the measurement of reading speed as a function of print size at 10° in the inferior visual field. Oral reading speeds for single sentences were measured using the rapid serial visual presentation (RSVP) paradigm. Stimuli were generated and presented using an SGI O2 workstation (Silicon Graphics Inc.) and a Sony color graphics display monitor (Model# GDM-17E21, refresh rate = 75 Hz) controlled by custom-written software. Words were rendered in Times-Roman, and were presented as high contrast (ca. 90%) black letters on a white background of 45 cd/m^2, at 10° below a horizontal fixation line (Chung, 2002; Chung et al., 1998). For all observers except SW, six print sizes, ranging from 0.56° to 3.2°, were tested. Observer SW was tested using print sizes ranging from 0.43° to 2.3°. Procedures and sentences were identical to those used by Chung et al. (1998) and Chung (2002). In brief, we used the Method of Constant Stimuli to present words in a sentence, one at a time, in a rapid sequence. Each word was exposed for a fixed duration. Only one print size was tested in each block of 18 trials (six word exposure durations with three trials per duration). The order of testing the six print sizes was randomized in the first half of each session, and was then reversed in the second half of the session. In other words, there were altogether 12 blocks of trials (six print sizes each tested twice) tested in the session. Observers were asked to read each sentence as quickly and as accurately as possible while fixating the fixation line. The number of words read correctly was recorded for each trial. None of the observers read any of the sentences more than once. Although we did not use an eye-tracker, we monitored observers' fixation by looking, from the side, the gaze of the observers for vertical eye movements. Horizontal eye movements along the fixation line were allowed, although most observers preferred to fixate at a certain point on the fixation line rather than scanning horizontally along the fixation line. A trial was discarded and repeated with a different sentence when vertical eye movements were detected. Averaged across observers, approximately 10% of trials were discarded and repeated. This trial rejection rate was similar to that reported by Chung et al. (1998) where an eye-tracker was used to monitor observers' fixation.



Pre-tests:
1. Reading speeds for six print sizes
2. Performance for identifying flanked letters at five letter separations (plus the unflanked condition)

Training:
60 blocks (10 blocks per day) of identifying flanked letters at the smallest letter separation

Post-tests:
1. Performance for identifying flanked letters at five letter separations (plus the unflanked condition)
2. Reading speeds for six print sizes

Fig. 1. A schematic cartoon illustrating the basic experimental design of the study.

Data obtained for each print size were analyzed as follows. We used a cumulative-Gaussian function to relate the proportion of words read correctly as a function of word exposure duration, from which we derived our criterion reading speed based on the RSVP exposure duration that yielded 80% of words read correctly, as in our previous studies (e.g., Chung, 2002; Chung et al., 1998, 2004; Yu et al., 2007). By plotting the criterion reading speed as a function of print size, we could extract two important parameters of reading performance (Fig. 5): maximum reading speed and critical print size (the smallest print size at which maximum reading speed could still be attained). The critical print size obtained in the pre-test session was then used to specify the letter size used in subsequent testing on letter identification performance.

2.2. Letter identification measurements

Performance for identifying the middle flanked letters in trigrams at 10° in the inferior visual field was measured as a function of letter separation during pre- and post-tests. Stimuli were generated on a Macintosh G4 computer with software written in Matlab (The MathWorks, MA), using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997), and were displayed on a Mitsubishi Diamond Pro 15'' monitor (model number N0701) at a background luminance of 40 cd/m^2. On each trial, a trigram consisted of three lowercase letters independently and randomly drawn from the 26 letters of the Times-Roman alphabet was presented at 10° below a small fixation target. Observers were asked to carefully fixate the fixation target before and during the presentation of each trigram. Trigrams were presented for 150 ms, a duration shorter than the latency of saccadic eye movements. The task of the observers was to identify the middle flanked letter of each trigram while fixating at the fixation target, and indicated their responses using a keyboard. Audio feedback was provided to indicate whether or not the response was correct. Although not an inclusion criterion for the study, all our observers were proficient computer users who typed well without looking at the keyboard; consequently the rate of finger errors was very low, as reflected in their perfect or near-perfect performance for the unflanked condition and at large letter separations following training (see Fig. 4).

To ensure that letter size was not a limiting factor, we used a letter size corresponding to 1.4× the critical print size as determined from the first pre-test session. Previous studies have shown that the spatial extent of crowding in the periphery does not depend on target size once the target size is above recognition level (Levi, Hariharan, & Klein, 2002a; Pelli, Palomares, & Majaj, 2004; Tripathy & Cavanagh, 2002), therefore we believe that our results would generalize to other letter sizes. Letter separations (center-to-center) were specified as multiples of the letter size (*x*-height) and included 0.8×, 1×, 1.25×, 1.6× and 2× (see Fig. 2 for sample trigrams). In addition, performance for identifying unflanked (single) letters was also measured. Each letter separation was tested in a separate block of 20 trials. The order of testing the six letter separations (including the unflanked condition) was randomized for each observer. As for the reading task, we used Times-Roman font for the task of letter identification. Although Times-Roman is a proportional-width font, we rendered the trigrams such that the center-to-center separation between adjacent letters was a fixed distance, when expressed as a multiple of *x*-height. At the smallest letter separation (0.8×), adjacent letters frequently touched but did not significantly overlap with one another, except for the wider letters 'w' and 'm'. How-

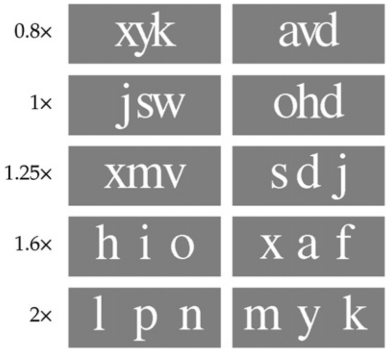


Fig. 2. Samples of trigrams rendered at different letter separations, specified as the distance between centers of adjacent letters and expressed as multiples of *x*-height (shown on the left). Each row shows two sample trigrams rendered at the specified letter separation.

ever, each letter was randomly drawn from the entire set of 26 letters with equal probability. As such, the odds of having a trigram comprised of only 'w' and 'm' was low.

During training, letter identification performance was only assessed at the closest letter separation (0.8×). Each block consisted of 100 trials, otherwise details of the measurements were identical to those described above.

3. Results

Performance for identifying a letter closely flanked by two other letters improved following six days of training for all observers. This improvement can be observed in our data in two ways. Averaged across observers, proportion-correct of letter identification increased from 0.34 at pre-tests to 0.64 at post-tests, representing a statistically significant improvement of 88% (paired *t*-test: $t_{(\mathrm{df}=7)} = 22.4$, $p < .0001$). Further, to track the changes in performance overtime, we plotted the accuracy of letter identification as a function of training block for each observer in Fig. 3. It is clear that performance improved with time (training block) for all observers. To quantify the improvement, we fit a linear regression function to each set of data that included measurements obtained for all training blocks. A *t*-test was performed to determine if the slope of each regression function differed from a slope of 0, an indication that there was no improvement due to training. Consistent across all observers, the *p*-value of the *t*-test was <.0001, implying that the improvement was statistically significant. Note that even if we applied the conservative Bonferroni correction to correct for multiple comparisons, the *p*-values for the slopes of all the regression functions remain statistically significant. These analyses provide clear evidence that performance for identifying crowded letters improved with training, a phenomenon to which we shall refer as "uncrowd".[1]

A signature of crowding is that the performance of identifying a target improves with the distance between the target and its flankers (Bouma, 1970). As shown in Fig. 4, proportion correct of identifying a letter in the presence of flanking letters improves with letter separation for all observers. However, the interesting question is whether or not the improvement following training in identifying crowded letters is specific to the trained letter separation, or whether it transfers to other letter separations. Fig. 4 clearly shows that the improvement in letter identification performance occurred not only at the trained letter separation (represented by the arrows close to the *x*-axes), but also at all other letter separations. The magnitudes of improvement at the untrained letter separations are not necessarily the same as that obtained at the trained separation, especially at the largest letter separation, probably because of a ceiling effect in performance.

Crowding is often attributed to a pooling (integration) of signals over space (e.g., Levi, Klein, & Hariharan, 2002b; Levi et al., 2002a; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pelli et al., 2004). To test if the spatial extent (integration zone) of crowding changes with training, we estimated the spatial extent of crowding following the method and definition of Kooi, Toet, Tripathy, and Levi (1994). We fit each set of proportion-correct vs. letter separation data as shown in Fig. 4 with a cumulative-Gaussian function, and determined the letter separation that corresponded to 50% correct of letter identification (after correction for guessing). Table 1 summarizes the spatial extent obtained before and after training for each observer. Averaged across observers, the spatial extent decreased from 1.12× letter size at pre-tests to 0.69× letter size at post-tests, representing a 38% reduction in the extent (paired *t*-test: $t_{(\mathrm{df}=7)} = 7.72$, $p = .0001$).

So far, we showed that performance for identifying crowded letters improves with training, and the improvements also transfer to other letter separations. But, does the improvement lead to higher reading speeds? Fig. 5 compares reading speeds for different print sizes, before and after training. Clearly, the reading speed vs. print size plots are not distinctively different between the pre- and post-test results. To quantify the reading performance, we fit each set of reading speed vs. print size data using a two-line fit on log–log axes (Chung, 2002; Chung et al., 1998, 2004), where the intersection of the two lines represents the critical print size (CPS), the smallest print size at which maximum reading speed could still be attained (Mansfield, Legge, & Bane, 1996). The slopes (on log–log axes) of the first and the second line were constrained to 2.32 and zero, respectively, based on the empirical finding of Chung et al. (1998) showing that the slope of the first line did not vary systematically with eccentricity, and averaged 2.32 across all observers and eccentricities up to at least 20°. Table 1 summarizes the log maximum reading speed and the critical print size before and after training for each observer. Averaged across observers, log maximum reading speed changed from 2.29 (194.98 wpm) at pre-tests to 2.32

---

[1] Crowding refers to the degradation of performance in the presence of nearby flankers. Here, we examined if the degradation of performance could be reduced or eliminated through training. We refer to the improvement in performance of identifying crowded letters as "uncrowd", which represents a release of the crowding effect. Because "uncrowd" is not a real word, it appears in this paper in quotes.

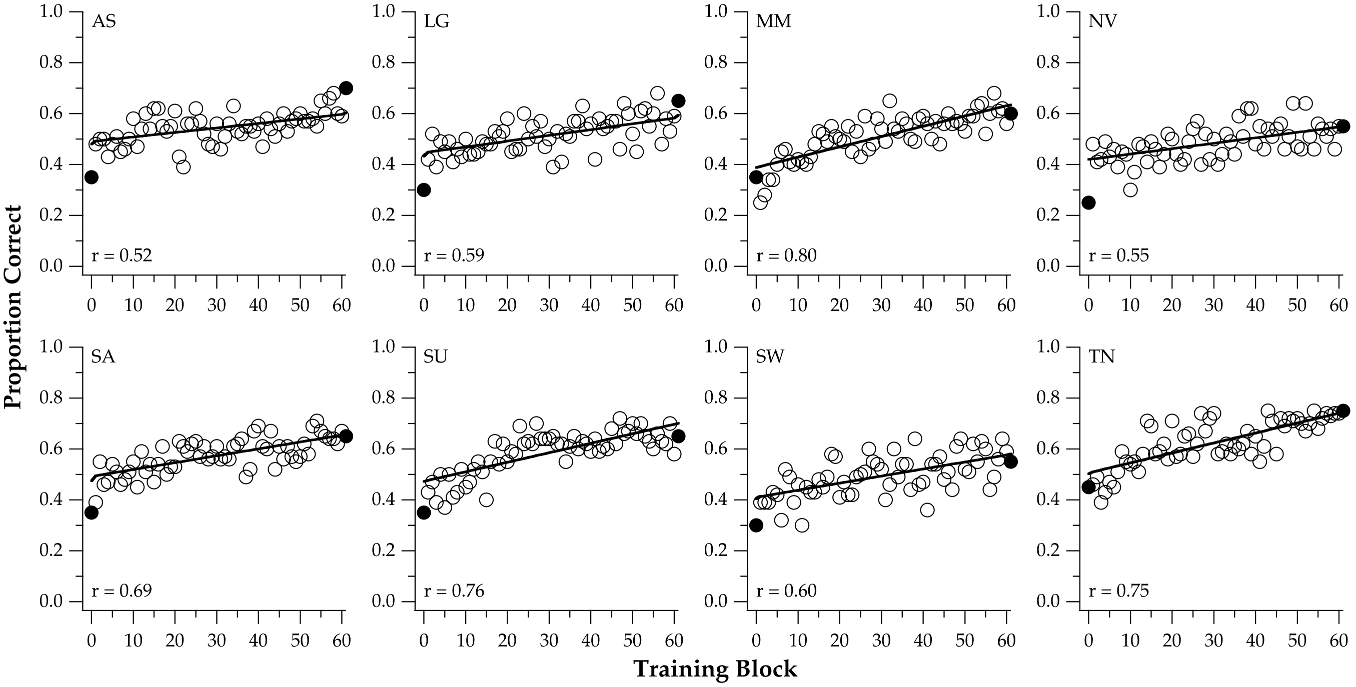


Fig. 3. Proportion-correct of identifying a letter flanked by two other letters at the closest letter separation (0.8× the *x*-height) is plotted as a function of training blocks, for each individual observer. Filled symbols in each panel represent measurements obtained at pre- and post-tests (not included in the fitting of the regression line). The solid line in each panel represents the best-fit regression line to the 60 blocks of training data. The correlation coefficient of the line is given in each panel. For all observers, the slope of the line is significantly different from 0 ($p < .0001$), implying a significant amount of improvement following training.

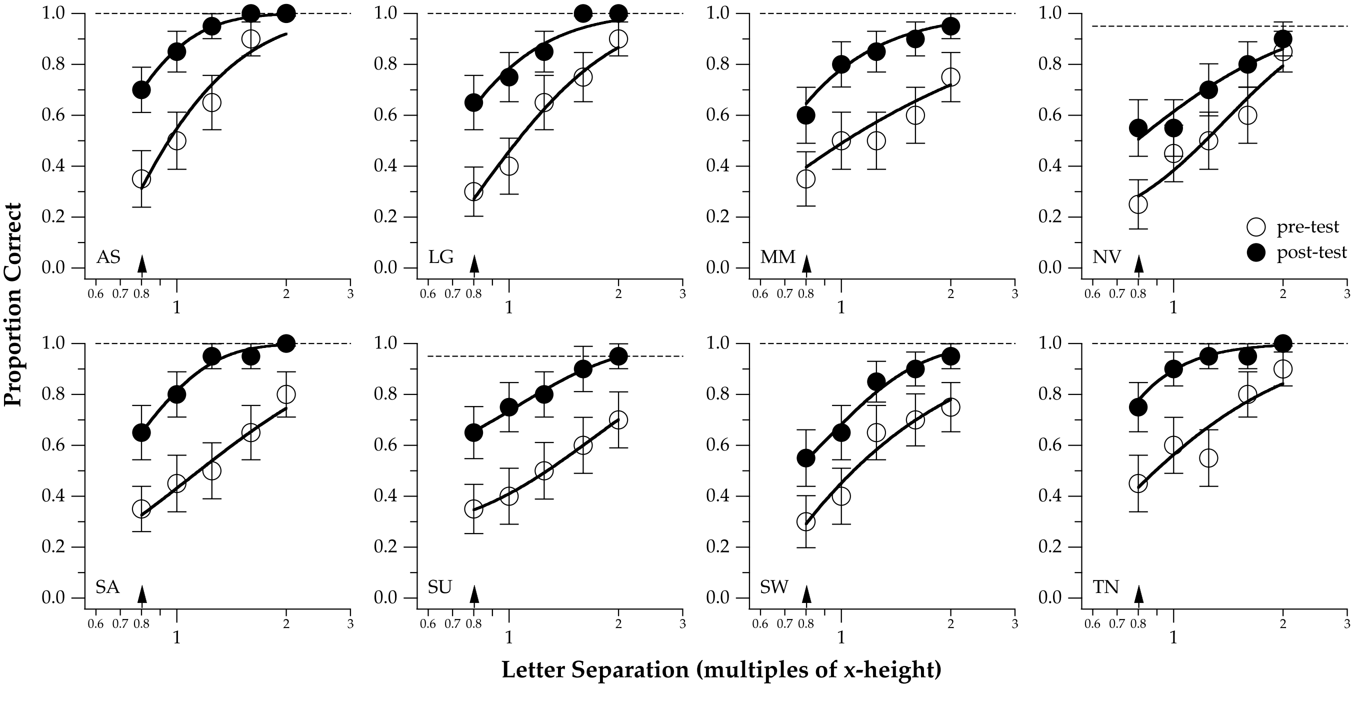


Fig. 4. Proportion-correct of identifying the middle letter of trigram is plotted as a function of letter separation (multiples of *x*-height) for each observer. Data for pre- and post-tests are represented by unfilled and filled symbols, respectively. Each set of data was fit with a cumulative-Gaussian function. Error bars represent ± 1 standard error of the proportion. The spatial extent of crowding is defined as the letter separation that yields 50% correct of letter identification (after correction for guessing) on the cumulative-Gaussian function. Dashed lines represent the performance for identifying unflanked (single) letters at pre-test (performance at post-test was either identical to pre-test or better). Arrows indicate the trained letter separation (0.8× the *x*-height).

Table 1
Comparisons of pre- and post-test performance for the spatial extent of crowding, maximum reading speed and critical print size for each observer

Observer	Spatial extent of crowding (multiples of x-height)		Log maximum reading speed		Critical print size (°)	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
AS	0.97	0.67	2.28	2.31	0.95	0.88
LG	1.08	0.71	2.30	2.39	1.38	1.34
MM	1.08	0.69	2.28	2.29	1.64	1.27
NV	1.26	0.82	2.28	2.27	1.40	1.34
SA	1.20	0.68	2.42	2.43	1.26	1.08
SU	1.33	0.56	2.13	2.11	1.50	1.30
SW	1.12	0.77	2.30	2.34	0.97	0.94
TN	0.93	0.63	2.36	2.39	1.17	0.95
Average	1.12	0.69	2.29	2.32	1.28	1.14

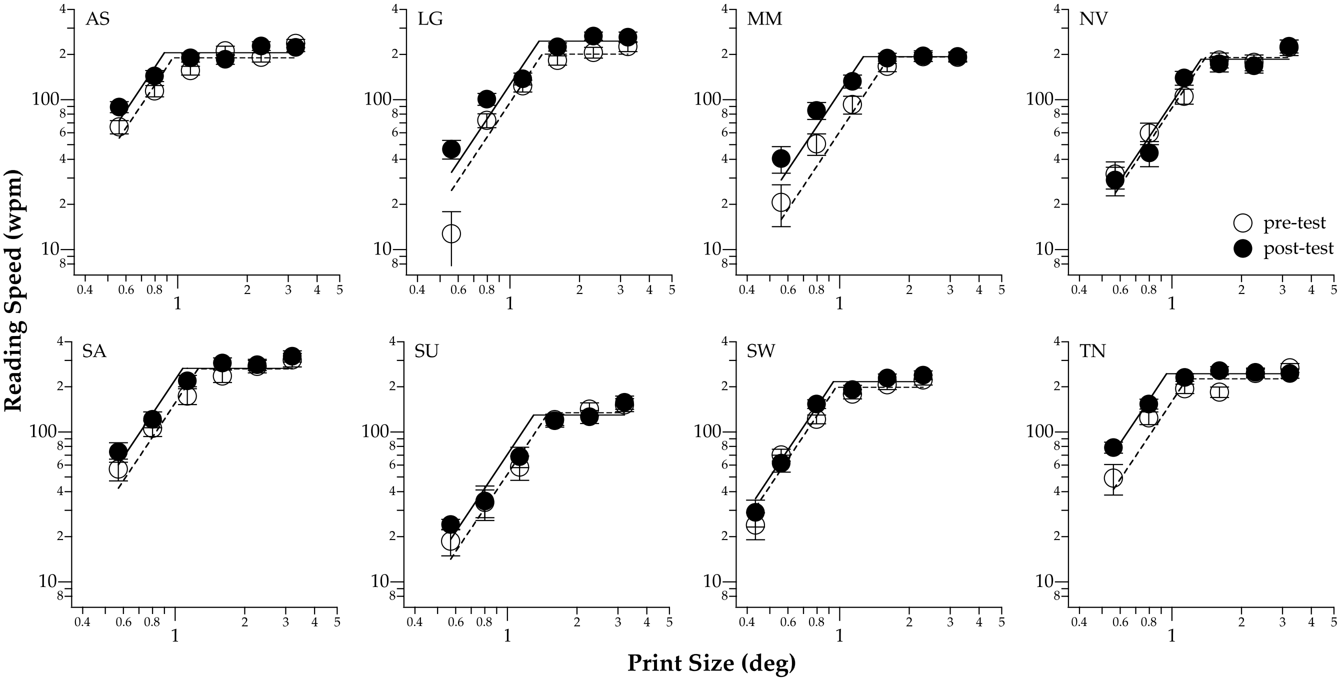


Fig. 5. RSVP reading speed (wpm) is plotted as a function of print size (deg) for each observer. Data for pre- and post-tests are represented by unfilled and filled symbols, respectively. Each set of data was fit with a two-line fit on log–log axes (see text for details) from which the maximum reading speed and critical print size were determined. Error bars represent ± 1 SEM.

(208.93 wpm, a 7.2% improvement) at post-tests, which are not statistically different from one another (paired t-test: $t_{(\mathrm{df}=7)} = -1.79$, $p = .12$). In contrast, critical print size for reading decreased from 1.28° at pre-tests to 1.14° at post-tests. This reduction in the critical print size is statistically significant (paired t-test: $t_{(\mathrm{df}=7)} = 3.51$, $p = .01$). Practically, our findings imply that even though our observers did not read faster following training to "uncrowd", they could read smaller print at the same maximum reading speed.

4. Discussion

Following six days (6000 trials) of repeated training on the task of identifying crowded letters at 10° inferior visual field, observers' performance for identifying such letters improved substantially. This improvement following training transferred to other untrained letter separations such that the spatial extent of crowding, defined as the smallest separation that a target letter needs to be separated from its flanking letters to yield an identification accuracy of 50%, decreased with training. In other words, the adverse effect of crowding can be reduced through training in terms of both the magnitude and extent of the effect. Unfortunately, the improvement in letter identification performance did not lead to faster reading speed, contrary to what one would predict if crowding is the primary limiting factor on reading speed in peripheral vision.

The finding that letter identification performance improved following training may not be surprising given that we have previously shown that performance for identifying single letters (Chung, Levi, & Tjan, 2005) or groups of three letters (Chung et al., 2004) improved with practice. However, the task and thus, presumably what observers learned through training, was different between this study and those of our previous studies. In this study, observers had to identify only one letter, but one that was flanked by two other letters in close proximity to the target letter. Therefore, the task differed from those in our previous two studies in that here, observers were faced with three letters on each trial, but that they had to learn to ignore or suppress the information from the two flanking letters that were irrelevant, and may even be distracting, to the task of identifying the middle flanked letters. Can our finding help us better understand the underlying mechanism of crowding?

To date, the underlying cause of crowding remains controversial. Currently, the two most popular theories for crowding postulate that crowding arises as a result of (1) an insufficient resolution of the attention system (He, Cavanagh, & Intriligator, 1996; Intriligator & Cavanagh, 2001; Leat, Li, & Epp, 1999; Strasburger, 2005; Strasburger, Harvey, & Rentschler, 1991; Tripathy & Cavanagh, 2002) or (2) an inherent deficiency in the bottom-up feature-integration process (Levi et al., 2002a, 2002b; Parkes et al., 2001; Pelli et al., 2004). With respect to the attention theory, our finding that training improves identification performance for crowded letters can be simply explained as a result of our observers being more capable of focusing their attention toward the flanked letters instead of dispersing their attention over the flankers. Alternatively, with respect to the feature-integration theory, the improvement in letter identification performance can be explained as a consequence of our observers being more capable of integrating the correct letter features from the flanked letter, instead of integrating the incorrect letter features from the flankers. In other words, our finding is consistent with the prediction of both theories. However, testing and differentiating between the viable theories of crowding was not the goal of the present study.

4.1. Present study vs. Chung et al. (2004)

Previously, we have shown that an improvement in letter identification performance following training in peripheral vision led to a 41% improvement in peripheral reading speed (Chung et al., 2004), a result that seems to be at odds with our current finding of a mere 7.2% improvement in the maximum reading speed. In both studies, reading speeds were measured using identical methods and the same number of trials. There were six observers in each training group in Chung et al. (2004) whereas we had eight observers in the present study, therefore, our failure to find an improvement in reading speed following training cannot be attributed to a lack of statistical power in the present study. We believe that the seemingly contradictory results between the Chung et al's (2004) study and the present study may in fact help us better understand the relationship among letter identification, crowding and reading. The goal of our 2004 study was to determine whether or not the visual span could be enlarged through training. Consequently, the training task was also the task that was used to measure the visual span, i.e., a trigram identification task that requires observers to respond to all three letters of a trigram. In the present study, our goal was to examine if observers could learn to "uncrowd", i.e., to minimize the detrimental effect of crowding. Traditionally, performance in the presence of crowding is assessed by asking observers to report the identity, or the orientation, or some other attributes of the target that is being crowded, but not those of the flankers (Chung, Levi, & Legge, 2001; Levi, Klein, & Aitsebaomo, 1985; Pelli et al., 2004). Here, we followed tradition and only required observers to identify the middle letters. The differences in task between the present study and our 2004 study could be crucial in at least two aspects. First, in the present study, we only measured and trained letter identification performance at the same letter position, *viz*, the letter position 10° directly below the fixation target. Even though we varied the letter separation within the trigram, the middle letter of each trigram occupied the exact same location. In contrast, in our earlier study, we trained observers at many letter positions right and left of the midline. Given that the majority of words comprise more than one letter, it is possible that an improvement in reading speed would only occur if there are improvements in letter identification performance across *all* letter slots, instead of an improvement at only one single slot. However, to explain the seemingly contradictory result between the present study and our earlier study, we have to further assume that the improvement following training to "uncrowd" at one letter position does not transfer to other letter positions.[2] Whether this assumption is true or not would need to be verified in future studies.

Second, the task of identifying only the middle letters of trigrams (albeit this is the conventional task for studying crowding, see above) might have encouraged observers to ignore the two flanking letters; whereas in our earlier (2004) study, observers were required to identify all three letters, hence they were not trained to disregard the two flanking letters. As mentioned earlier, the majority of words comprise more than one letter. A word can only be correctly identified if there is letter identity information from more than one single letter position. In other words, even if we have perfect identification performance (100% correct) for one single letter of a word but do not have information on its neighboring letters, the odds are that we will not be able to identify the word. However, given that we did not measure the identification performance of

---

[2] In the study of Chung et al. (2004), the increase in the size of the visual span (expressed as bits of information transmitted) following training transferred between the upper and lower visual fields (at the same eccentricity). However, given that the location specificity of the improvements following perceptual learning seems to be task-specific (e.g., Beard, Levi, & Reich, 1995; Fiorentini & Berardi, 1980; Fiorentini & Berardi, 1981; Kapadia, Gilbert, & Westheimer, 1994; Karni & Sagi, 1991; Sireteanu & Rettenbach, 2000), it is unclear whether or not the improvement following training to "uncrowd" transfers to other untrained letter positions.

the two flankers, we did not know if the performance for identifying the two flankers improved (as a result of a transfer of learning to other letter position) or declined (if observers learned to suppress the information from the flankers), or remained the same following training. Consequently, whether the lack of an improvement in reading speed following training was due to the suppression of information from the two flankers would need to be investigated in future studies.

4.2. Reading speed improvements at small print sizes?

Although it is clear from Fig. 5 and Table 1 that maximum reading speeds of our observers did not benefit from our training task, all observers except NV and SW showed a small improvement in reading speed for print sizes smaller than the critical print size. This observation is consistent with the reduction in critical print size following training that is not associated with a change in the maximum reading speed. The magnitude of crowding has been shown to be stronger for small than for large print (Arditi, Knoblauch, & Grunwald, 1990; Chung, 2002; Yu et al., 2007). Specifically, reading speed was found to benefit from a slight increase in letter spacing in text, for spacing that is smaller than standard spacing in printed text (Chung, 2002; Yu et al., 2007). The benefit was more obvious for small than for large print sizes. Based on this finding, one might speculate that the improvement following learning to “uncrowd” would lead to an improvement in reading speed only at smaller print sizes, where crowding is presumably stronger. To determine whether or not this speculation was correct, we plotted in Fig. 6 the ratios of post-test and pre-test reading speeds (a value >1 signifies improvement) as a function of normalized print sizes—print sizes normalized to each observer’s own critical print size. Note that this comparison of post/pre reading speed for each print size per observer is noisy, that is why we used curve-fitting to summarize each observer’s reading performance earlier (Fig. 5). However, here, we are interested in finding if there exists a trend showing improvement in reading speed following training for smaller print sizes.

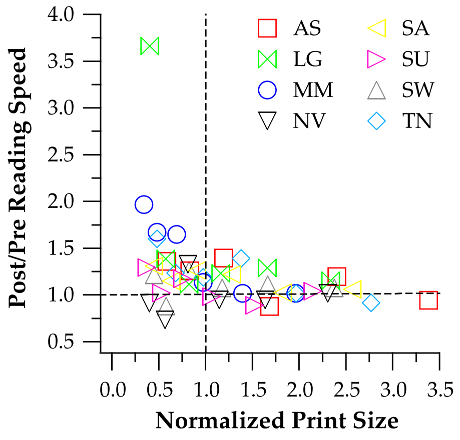


Fig. 6. Ratios of post-test and pre-test reading speed were plotted as a function of individual print sizes normalized to each observer’s critical print size. The vertical dashed line divides the print size axis into those smaller (left of the line) or larger (right of the line) than the critical print size. The horizontal dashed line divides the reading speed ratio axis into those showing an improvement following training (ratios >1) and those showing a decline in performance following training (ratios <1).

Given that the critical print size represents the smallest print size an observer requires in order to read at the maximum reading speed, we adopt it as the criterion to separate print sizes into “small” (those smaller than the critical print size) and “large” (those larger than the critical print size). Fig. 6 shows that for the “large” print sizes, most of the post/pre reading speed ratios fell very close to the line of 1 (no improvement), with an occasional few points falling between the ratio of 1–1.4. For the “small” print sizes, there seems to be more individual variability. While two observers (NV and SW) did not seem to have any improvement (ratios close to or even less than 1) in reading speed following training, two other observers (LG and MM) showed clear improvement (ratios >1) in reading speed following training, with the rest of the observers yielding a ratio between 1 and 1.5. Overall, there seems to be a trend that there is some improvement in reading speed following our training task at small print sizes, but the magnitude of the improvement is obscured by the intrinsic variability associated with the comparison of post/pre reading speed for each print size, as well as individual observer variability.

4.3. Spatial extent of crowding shrinks following training

In this study, we found that the spatial extent of crowding decreases following learning to identify letters closely flanked by other letters. What accounts for the shrinkage in the spatial extent of crowding following training to “uncrowd”? The spatial extent of crowding has been discussed in relation to the receptive field size of neurons that presumably underlie crowding (e.g., Chung, Li, & Levi, 2007; Pelli et al., 2004; Toet & Levi, 1992). If this were true, then we would expect that the receptive field size of neurons at the site at which crowding occurs would shrink in size following perceptual learning. Unfortunately, physiological evidence directly linking receptive field size, crowding and perceptual learning does not exist. This could be attributed to, in part, the fact that we are still unclear as to where the neural origin of crowding is. Area V4 has been implicated as the site for crowding in macaque monkey (Motter, 2002). More recently, using the functional magnetic resonance imaging (fMRI) technique, we found that the blood oxygenation level dependent (BOLD) response amplitude for the signal, relative to the fixation baseline condition, was significantly lower when human observers were asked to identify crowded letters, compared with single letters (non-crowded condition), at visual area V2 and beyond. In contrast, there was no difference in the BOLD response amplitude between the crowded and non-crowded condition in V1. This result rules out V1 as the cortical site for crowding and places the site for crowding at visual area as early as V2 (Arman, Chung, & Tjan, 2006).

If V2 is indeed the site for crowding, is there evidence to suggest that the receptive field properties of neurons in V2 could be modified through perceptual learning? Following extensive training on an orientation discrimination task, Ghose, Yang, and Maunsell (2002) failed to find any changes in the selectivity or responsiveness of neurons in V1 or V2 that could account for the improvement in performance in their monkeys, although a later study from the same laboratory showed that basic neuronal response properties in V4 were modified following training (Yang & Maunsell, 2004). In contrast, by training adult monkeys reared with abnormal visual experience on spatial contrast sensitivity and stereoacuity tests, Nakatsuka et al. (2007) found that disparity sensitivity was significantly better in trained than untrained monkeys in V2, but not in V1. The result of Nakatsuka et al. (2007) provides evidence that neurons in V2 are indeed malleable in response to perceptual learning and could exhibit neuronal changes following perceptual learning. Clearly, whether or not the spatial extent of crowding indeed directly relates to neuronal receptive field size, specifically, receptive fields at V2, would need to be further investigated.

5. Conclusion

In this study, we succeeded in reducing crowding by training observers to identify crowded letters in peripheral vision (“uncrowd”). Because the improvement in letter identification performance improved at the smallest, trained letter separation, as well as all other untrained but larger letter separations, we conclude that our training paradigm is effective in reducing both the magnitude and the spatial extent of crowding. Unfortunately, we failed to find an accompanied improvement in the maximum reading speed, suggesting that crowding is unlikely to be the primary limiting factor on maximum reading speed in peripheral vision.

In addition, our quest for using a simpler training paradigm than the one used in Chung et al. (2004) to obtain the same magnitude of improvement in reading speed failed. However, this failure could be due to the differences in the tasks between the present study and Chung et al. (2004). Currently, we are investigating whether reading speed can be improved by controlling each of these differences. We are also exploring other training paradigms that may be effective in enhancing peripheral reading speed.

Acknowledgments

This study was supported by NIH Grant R01-EY12810. I thank Drs. Gordon Legge and Bosco Tjan for helpful discussion.

References

Arditi, A., Knoblauch, K., & Grunwald, I. (1990). Reading with fixed and variable character pitch. *Journal of the Optical Society of America A, 7*, 2011–2015.

Arman, A. C., Chung, S. T. L., & Tjan, B. S. (2006). Neural correlates of letter crowding in the periphery [Abstract]. *Journal of Vision, 6*(6), 804, 804a.

Beard, B. L., Levi, D. M., & Reich, L. N. (1995). Perceptual learning in parafoveal vision. *Vision Research, 35*, 1679–1690.

Berger, T. D., Martelli, M., Su, M., Aguayo, M., & Pelli, D. G. (2003). Reading quickly in the periphery [Abstract]. *Journal of Vision, 3*(9), 806, 806a.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*, 177–178.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Bullimore, M. A., & Bailey, I. L. (1995). Reading and eye movements in age-related maculopathy. *Optometry and Vision Science, 72*, 125–138.

Chung, S. T. L. (2002). The effect of letter spacing on reading speed in central and peripheral vision. *Investigative Ophthalmology and Visual Science, 43*, 1270–1276.

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*, 695–709.

Chung, S. T. L., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research, 41*, 1833–1850.

Chung, S. T. L., Levi, D. M., & Tjan, B. S. (2005). Learning letter identification in peripheral vision. *Vision Research, 45*, 1399–1412.

Chung, S. T. L., Li, R. W., & Levi, D. M. (2007). Crowding between first- and second-order letter stimuli in normal foveal and peripheral vision. *Journal of Vision, 7*(2), 10, 1–13.

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*, 2949–2962.

Cline, D., Hofstetter, H. W., & Griffin, J. R. (1997). *Dictionary of visual science* (4th ed.). Boston: Butterworth-Heinemann.

Elliott, D. B., Trukolo-Ilic, M., Strong, J. G., Pace, R., Plotkin, A., & Bevers, P. (1997). Demographic characteristics of the vision-disabled elderly. *Investigative Ophthalmology & Visual Science, 38*, 2566–2575.

Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature, 287*, 43–44.

Fiorentini, A., & Berardi, N. (1981). Learning in grating waveform discrimination: Specificity for orientation and spatial frequency. *Vision Research, 21*, 1149–1158.

Flom, M. C. (1991). Contour interaction and the crowding effect. *Problems in Optometry, 3*, 237–257.

Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963). Visual resolution and contour interaction. *Journal of the Optical Society of America, 53*, 1026–1032.

Ghose, G. H., Yang, T., & Maunsell, J. H. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. *Journal of Neurophysiology, 87*, 1867–1888.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*, 334–337.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43*, 171–216.

Jacobs, R. J. (1979). Visual resolution and contour interaction in the fovea and periphery. *Vision Research, 19*, 1187–1195.

Kapadia, M. K., Gilbert, C. D., & Westheimer, G. (1994). A quantitative measure for short-term cortical plasticity in human vision. *Journal of Neuroscience, 14*, 451–457.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Science of United States of America, 88*, 4966–4970.

Kleen, S. R., & Levoy, R. J. (1981). Low-vision care: Correlation of patient age, visual goals, and aids prescribed. *American Journal Optometry and Physiological Optics, 58*, 200–205.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision, 8*, 255–279.

Latham, K., & Whitaker, D. (1996). A comparison of word recognition and reading performance in foveal and peripheral vision. *Vision Research, 36*, 2665–2674.

Leat, S. J., Li, W., & Epp, K. (1999). Crowding in central and eccentric vision: The effects of contour interaction and attention. *Investigative Ophthalmology & Visual Science, 40*, 504–512.

Leat, S. J., & Rumney, N. J. (1990). The experience of a university-based low-vision clinic. *Ophthalmic and Physiological Optics, 10*, 8–15.

Legge, G. E., Cheung, S.-H., Yu, D., Chung, S. T. L., Lee, H.-W., Owens, D. P. (2007). The case for the visual span as a sensory bottleneck in reading. *Journal of Vision, 7*(2), 9, 1–15.

Legge, G. E., Mansfield, J. S., & Chung, S. T. L. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research, 41*, 725–743.

Levi, D. M., Hariharan, S., & Klein, S. A. (2002a). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision, 2*, 167–177.

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research, 25*, 963–977.

Levi, D. M., Klein, S. A., & Hariharan, S. (2002b). Suppressive and facilitatory spatial interactions in foveal vision: Foveal crowding is simple contrast masking. *Journal of Vision, 2*, 140–166.

Levi, D. M., Song, S., & Pelli, D. G. (2007). Amblyopic reading is crowded. *Journal of Vision*, in press.

Loomis, J. M. (1978). Lateral masking in foveal and eccentric vision. *Vision Research, 18*, 335–338.

Manny, R. E., Fern, K. D., Loshin, D. S., & Martinez, A. T. (1988). The effects of practice on contour interaction. *Clinical Vision Science, 3*, 59–67.

Mansfield, J. S., Legge, G. E., & Bane, M. C. (1996). Psychophysics of reading – XV: Font effects in normal and low vision. *Investigative Ophthalmology & Visual Science, 37*, 1492–1501.

Motter, B. C. (2002). Crowding and object integration within the receptive field of V4 neurons [Abstract]. *Journal of Vision, 2*(7), 274, 274a.

Nakatsuka, C., Zhang, B., Watanabe, I., Zheng, J., Bi, H., Ganz, L., et al. (2007). Effects of perceptual learning on local stereopsis and neuronal responses of V1 and V2 in prism-reared monkeys. *Journal of Neurophysiology, 97*, 2612–2626.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4*, 739–744.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 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*, 1136–1169.

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*, in press.

Rubin, G. S., & Turano, K. (1994). Low vision reading with sequential word presentation. *Vision Research, 34*, 1723–1733.

Sireteanu, R., & Rettenbach, R. (2000). Perceptual learning in visual search generalizes over tasks, locations, and eyes. *Vision Research, 40*, 2925–2949.

Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*, 1024–1037.

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

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research, 32*, 1349–1357.

Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research, 42*, 2357–2369.

Whittaker, S. G., Rohrkaste, F., & Higgins, K. (1989). Optimum letter spacing for word recognition in central and peripheral vision. In *Digest of topical meeting on noninvasive assessment of the visual system* (pp. 56–59). Washington, DC: OSA.

Yang, T., & Maunsell, J. H. (2004). The effect of perceptual learning on neuronal responses in monkey visual area V4. *Journal of Neuroscience, 24*, 1617–1626.

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.