Title

Learning to identify crowded letters: Does the learning depend on the frequency of training?

Authors
Chung, Susana TL
Truong, Sandy R

Publication Date
2013

DOI
10.1016/j.visres.2012.11.009

Peer reviewed


ELSEVIER
VISION
RESEARCH
An International Journal for Functional Aspects of Vision
L
R

Vision Research 77 (2013) 41–50

Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Learning to identify crowded letters: Does the learning depend on the frequency of training?

Susana T.L. Chung *, Sandy R. Truong

School of Optometry, University of California, Berkeley, Berkeley, CA 94720-2020, United States

A R T I C L E   I N F O

Article history:
Received 28 July 2012
Received in revised form 19 November 2012
Available online 30 November 2012

Keywords:
Perceptual learning
Crowding
Letter identification
Peripheral vision

A B S T R A C T

Performance for many visual tasks improves with training. The magnitude of improvement following training depends on the training task, number of trials per training session and the total amount of training. Does the magnitude of improvement also depend on the frequency of training sessions? In this study, we compared the learning effect for three groups of normally sighted observers who repeatedly practiced the task of identifying crowded letters in the periphery for six sessions (1000 trials per session), according to three different training schedules—one group received one session of training everyday, the second group received a training session once a week and the third group once every 2 weeks. Following six sessions of training, all observers improved in their performance of identifying crowded letters in the periphery. Most importantly, the magnitudes of improvement were similar across the three training groups. The improvement was accompanied by a reduction in the spatial extent of crowding, an increase in the size of visual span and a reduction in letter-size threshold. The magnitudes of these accompanied improvements were also similar across the three training groups. Our finding that the effectiveness of visual perceptual learning is similar for daily, weekly and biweekly training has significant implication for adopting perceptual learning as an option to improve visual functions for clinical patients.

1. Introduction

Performance for a variety of visual tasks improves with practice (e.g. Ball & Sekuler, 1982, 1987; Beard, Levi, & Reich, 1995; Fahle & Edelman, 1993; Fiorentini & Berardi, 1980, 1981; Karni & Sagi, 1991; McKee & Westheimer, 1978; Poggio, Fahle, & Edelman, 1992; Saarinen & Levi, 1995). The magnitude of improvement following the process of repeated practice (training), often termed *perceptual learning*, depends on many aspects of the training regime, including the task chosen for training, the total amount of practice and the amount of practice within each training session. Recently, perceptual learning has been proposed as a treatment to improve visual functions or to overcome some of the disabilities as a result of amblyopia (Astle, Webb, & McGraw, 2011; Levi & Li, 2009; Polat, 2009), presbyopia (Polat et al., 2012) and macular disorders (Chung, 2011). A major consideration for applying perceptual learning to improving vision in clinical patients is compliance, which usually relates to the inconveniences brought about by the training regime. For instance, if the training regime calls for many training sessions, or extensive hours of training for each session, patients may find it difficult to adhere to the training schedule. Fortunately, for many visual tasks, improvements usually occur fairly rapidly for the first couple of training sessions (e.g. Fiorentini & Berardi, 1981; Karni & Sagi, 1993; Poggio, Fahle, & Edelman, 1992), although it has been shown that performance for certain tasks could improve slowly after the initial rapid improvement; and may require up to 40–50 h of practice to reach a plateau (Li, Klein, & Levi, 2009; Li, Provost, & Levi, 2007). Also, shorter training sessions have been shown to be more effective in inducing improvements than longer ones (Molloy et al., 2012). Therefore, the number of training sessions and the duration of each session may not be the major factors limiting patient compliance.

Numerous studies that examined perceptual learning in observers with normal vision adopted a protocol in which observers attended daily training sessions (e.g. Chung, 2007; Chung, Legge, & Cheung, 2004; Chung, Levi, & Tjan, 2005), or at the minimum, three to five training sessions per week (e.g. Gold, Bennett, & Sekuler, 1999; Li, Klein, & Levi, 2009; Li, Provost, & Levi, 2007; Saarinen & Levi, 1995; Sun, Chung, & Tjan, 2010). This frequency of training sessions was believed to be crucial to maximize the benefit of perceptual learning. However, is daily training really necessary to obtain the largest amount of improvement? If perceptual learning is to be adopted as a treatment for clinical populations, relaxing the frequency of training sessions is necessary as many patients may not be able to attend daily training sessions. This is especially so for visually impaired patients who are not able

* Corresponding author.
E-mail addresses: s.chung@berkeley.edu (S.T.L. Chung), sandyrtruong@gmail.com (S.R. Truong).

0042-6989/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2012.11.009

to drive and thus their availability to attend training sessions would depend on arrangements for transportation. Therefore, the primary goal of this study was to examine the dependence of the efficacy of perceptual learning on the frequency of training sessions. We compared the amount of improvement following perceptual learning in the normal periphery for three training schedules: daily, weekly and biweekly (every 2 weeks). Daily training was the popular schedule used in many previous perceptual learning studies. We chose to evaluate the effectiveness of a weekly and biweekly training schedule because many visually impaired observers are able to participate in research projects in our laboratory once a week or once every 2 weeks, implying that a weekly or biweekly training schedule is feasible for this group of patients. However, it remains unknown whether the effectiveness of perceptual learning would be reduced when there is a longer time interval between training sessions.

Though virtually any training task could be used to address our primary goal, in this study, we trained normally sighted observers to identify letters closely flanked by two other letters (learning to "uncrowd") in the periphery. This training task has been proven to be effective in reducing crowding in the normal periphery (Chung, 2007; Sun, Chung, & Tjan, 2010) and in observers with amblyopia (Chung, Li, & Levi, 2012). Crowding refers to the deleterious influence of nearby contours on visual discrimination (Levi, 2008; Pelli, Palomares, & Majaj, 2004). Previously, we trained observers to identify a letter closely flanked by two other letters, and found that following 6000 trials of repeated testing, observers improved in their ability to identify the flanked letters. This effect was found in the normal periphery (Chung, 2007; Sun, Chung, & Tjan, 2010) as well as in the amblyopic eye of a group of amblyopic observers (Chung et al., 2012). Because our target and flanking letters were randomly chosen from the 26 lowercase letters on each trial, the observed improvements could not be attributed to observers learning a specific combination of letters, as in studies in which only a very limited set of combinations of letters was used for training (e.g. Huckauf & Nazir, 2007). Using a different paradigm in which the letter spacing between the flanking letters and the target letter varied during training, Hussain et al. (2012) reported a similar effect that crowding can be reduced in the normal periphery and in amblyopic observers through perceptual learning. Interestingly, even when the training task was not specifically designed to reduce crowding, such as in video-game playing (e.g. Green & Bavelier, 2007; Li et al., 2011), or using a task that is more closely related to lateral masking than crowding[1] (Maniglia et al., 2011), a reduction in crowding was still observed following perceptual learning. The reduction in crowding was manifested as either better acuities measured in the presence of flankers in close proximity, or a reduction in the target-flanker spacing such that the performance for discriminating some attribute of the target (e.g. contrast or orientation) was not affected (for a review, see Huurneman et al., 2012).

The design of this study closely followed that of Chung (2007) with some modifications. In the Chung (2007) study, despite a substantial improvement in observers' ability to identify crowded letters following a daily training protocol, the improvement did not lead to improved reading speed. Previously, Legge and coworkers showed that the visual span, the number of characters that can be recognized in a single glance, is a sensory bottleneck on reading (Legge, 2007; Legge et al., 2007). This supposition is based on the strong correlation ($r^2 > 0.8$) between reading speed and the size of the visual span (expressed as mutual information transmitted in bits, see Section 2) determined for different stimulus characteristics such as contrast, letter size and stimulus presentation eccentricity.

Given the link between reading speed and visual span, and the finding of Chung (2007), we expected that the visual span, like reading speed, would not benefit from the same uncrowd training task. Such a result would further strengthen the supposition of the visual span as a sensory bottleneck on reading. On the contrary, if the visual span benefits from the uncrowd training task, then the close relationship between the visual span and reading speed would need to be revisited, and the results might help us understand why reducing crowding does not benefit reading speed. The secondary goal of this study was to test if the improvements following a training protocol to learn to uncrowd would lead to an enlargement in the visual span.

To preview our results, we found that observers showed an improved ability to identify crowded letters following six sessions of training. Most importantly, the magnitudes of improvement were similar for the daily, weekly and biweekly training groups. The improvement due to training was accompanied by a reduction in the spatial extent of crowding, an increase in the size of the visual span and a reduction in letter-size threshold. The magnitudes of these (transferred) improvements were also similar among the three training groups.

2. Methods

Twenty-four young adults with normal vision, aged 19–27, participated in this study. Written informed consent was obtained from each observer after the procedures of the experiment were explained and prior to the commencement of data collection. Observers were randomly assigned to one of three training groups, with eight observers in each group.[2] The three training groups differed only on the frequency of the training sessions, with one group receiving training on a daily basis ("daily"), the second group received training on a weekly basis (once a week on the same day of the week, "weekly") and the third group received training every fortnight (once every other week on the same day of the week, "biweekly"). The average ages of the three groups were very similar (daily = 20.13 years, weekly = 20.75 years, biweekly = 20.5 years). All testings (pre-tests, training and post-tests) were performed at 10° eccentricity in the lower visual field.

The basic experimental design and training schedule are represented schematically in Fig. 1. The pre-test, lasted approximately 1.5 h, consisted of the measurements of letter-size threshold, spatial extent of crowding and a visual-span profile (in the order listed). The letter-size threshold was used to determine the letter size that was used in subsequent testings (other pre-tests and training).

Training consisted of six sessions, each lasting approximately 1 h. The training task was very similar to that used in Chung (2007), whereby observers identified a letter flanked closely by two other letters on each trial, at 10° in the inferior visual field (Fig. 2A). The only differences between this study and Chung (2007) were that we used Courier font in this study (Times font was used in Chung (2007)) and that we specified the letter separation with respect to the standard letter spacing (equivalent to

[1] The distinction between lateral masking and crowding has been addressed in previous studies (Chung, Levi, & Legge, 2001; Pelli et al., 2004).

[2] A power analysis for ANOVA designs revealed that our sample size of eight observers per group yielded a power of 0.999 to detect any effect at $p = 0.05$, for our training task of identifying crowded letters, as well as for the untrained tasks of spatial extent of crowding measurements and visual-span measurements. For the trained task of identifying crowded letters, we assumed an improvement in proportion-correct of 0.181, with a standard deviation of 0.048 (values based on finding of Chung (2007)). This yielded an effect size of 3.77. For the untrained tasks, the assumed effect sizes were 6.783 (average post-pre ratio = 0.624, standard deviation = 0.092, based on Chung (2007)) for the spatial extent of crowding measurements; and 3.414 (average improvements in bits = 6.1, standard deviation = 1.787, based on Chung, Legge, and Cheung (2004)) for visual-span measurements.

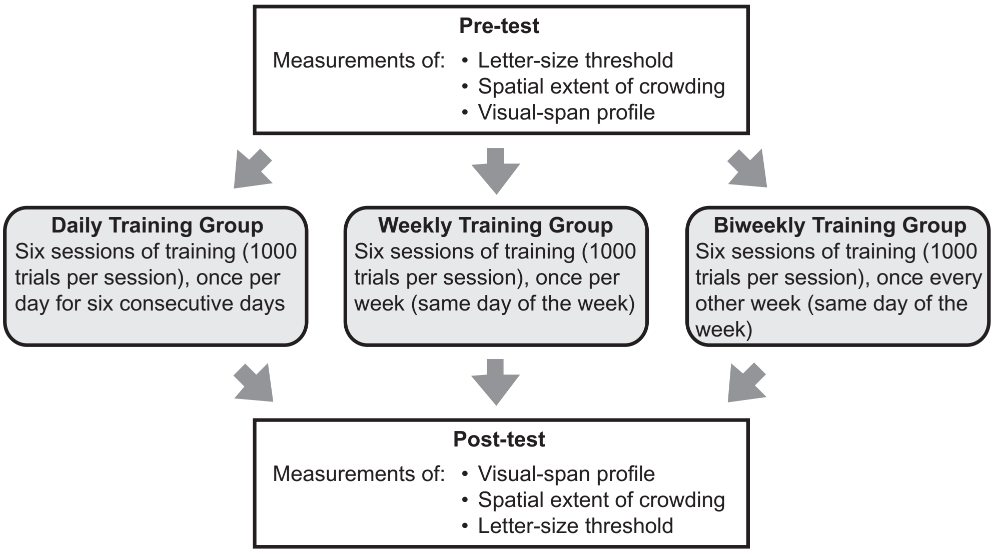

Fig. 1. A schematic cartoon illustrating the basic experimental design of the study.

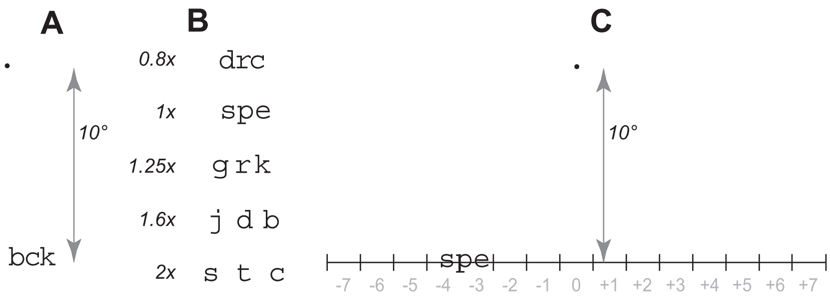

Fig. 2. Stimulus configurations for the different tasks. In (A), the trigram *bck* is presented at 10° directly below a small fixation dot. Observers' task was to identify the middle letter. Here, the center-to-center separation between letters is 0.8× the standard spacing, the separation that we used for training. For the pre- and post-test measurements of the spatial extent of crowding, the stimulus configuration was identical to that shown in (A), with the exception that we tested observers' performance for identifying the middle letter of trigrams for letter separations ranging from 0.8 to 2× the standard spacing. A sample trigram rendered at each of the letter separation is shown in (B). For the pre- and post-test measurements of the visual span, trigrams were rendered at the standard spacing and were presented at 13 letter positions (indexed by the middle letter of each trigram), from six letter slots left of, to six letter slots right of the vertical midline (C). Observers' task was to identify all three letters, from left to right. The gray lines and the numbers indicating letter positions are shown here for illustration purpose only. They were not shown on the monitor during testing.

1.16× the width of the lowercase letter *x*) in Courier. Each session comprised 10 blocks of trials, with 100 trials per block. Observers in the daily training group completed their training over six consecutive days, while observers in the weekly and biweekly training groups completed their training over six and eleven consecutive weeks, respectively.

The post-test immediately followed the last training session on the same day. Following the last training session, observers were given a 15–30 min break before the post-test commenced. The post-test was identical to the pre-test except that the measurements of the visual-span profile, spatial extent of crowding and letter-size thresholds were conducted in the reverse order as that during the pre-test.

2.1. Stimuli

Trigrams, random sequences of three letters arranged horizontally, were used as stimuli during all phases of the experiment (pre-test, training and post-test). With the exception of the visual-span measurements, the middle letter of each trigram was presented at 10° directly below a fixation target. For the visual-span measurements, the trigrams were presented at 10° below the fixation target, but at various letter positions right or left of the vertical midline (see details later). Stimuli were generated on a Macintosh G4 computer with software written in MATLAB 5.2.2 (The MathWorks, MA), using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997), and were displayed on a Sony color monitor (Model # GDM-17E21, refresh rate = 75 Hz). All trigrams comprised only lowercase letters rendered in Courier font, randomly chosen from the 26 letters of the Roman alphabet, and were rendered as black (0.2 cd/m^2) letters on a white background (45 cd/m^2). The center-to-center separation between adjacent letters of the trigrams varied, depending on the different task (see below). Each trigram was presented for 106 ms, a duration shorter than the latency of saccadic eye movements. The task of the observers was to identify only the middle flanked letter (all tasks except for the measurement of visual-span profiles) or all three letters (measurement of visual-span profiles) of each trigram while fixating the fixation target.

2.2. Pre- and post-test: Letter-size threshold measurements

Letter-size threshold was measured using letter trigrams, where adjacent letters were separated from each other by a center-to-center separation equivalent to 3× the standard letter spacing. This letter separation was large enough such that the measured size threshold approximated the unflanked threshold; while at the same time, maintained the same task demand for our observers.

We used the Method of Constant Stimuli to present trigrams at five letter sizes in each block. Observers' task was to identify the middle letter of each trigram. For each observer, a cumulative-Gaussian function was used to fit the data relating the proportion of correct responses and letter size. The letter-size threshold was defined as the letter size that yielded 52% correct (50% correct after correction for guessing) on the cumulative-Gaussian function.

2.3. Pre- and post-test: Spatial extent of crowding measurements

Performance for identifying the middle letters of trigrams was measured as a function of letter separation during pre- and post-tests, as in Chung (2007). Five letter separations were tested, ranging from 0.8 to 2× the standard letter spacing (Fig. 2B). Letter size was 1.4× the letter-size threshold as determined previously for each observer, which yielded a proportion of correct responses averaging 0.85 (range: 0.63–0.93) for the largest letter separation before training. A cumulative-Gaussian function was used to fit the data relating the proportion of correct responses and letter separation. The separation that yielded 52% correct (50% correct after correction for guessing) on the cumulative-Gaussian function was used to represent the spatial extent of crowding.

2.4. Pre- and post-test: Visual-span measurements

Visual-span profiles were measured using a trigram-recognition task as in Chung, Legge, and Cheung (2004). Letter size was fixed at 1.4× the letter-size threshold. For this task, trigrams were presented at 13 positions, indexed by the position of the middle letter, from six letter slots left of the vertical midline (letter slot 0 was 10° directly below fixation) to six letter slots right of the vertical midline (Fig. 2C). Each trigram position was tested ten times in a random order within a block of trials, yielding a total of 130 trials tested in each block. Observers' task was to identify all three letters of each trigram, from left to right. A letter was scored as being identified correctly if and only if its order within the trigram was also correct. To calculate the overall performance of letter identification at each letter slot, we combined the identification accuracies across trials where the letter slot was occupied by the left, middle or right letter of a trigram. A split-Gaussian function centered at letter slot 0 was then used to fit each set of data relating identification accuracy and letter position, representing the visual-span profile (Chung, Legge, & Cheung, 2004; Legge, Mansfield, & Chung, 2001). Curve-fitting was restricted to data within five letter slots left and right of fixation because the sixth letter slot left and right of fixation did not contain trials where the letter slot was occupied by the inner letter (the letter of a trigram closest to fixation). To quantify the size of the visual span, following Legge, Mansfield, and Chung (2001) and Chung, Legge, and Cheung (2004), we converted the identification accuracy at each letter slot to bits of mutual information transmitted by the visual span. According to Information Theory (Shannon, 1948), mutual information measures the amount of information that can be obtained about one random variable by observing another. In other words, it quantifies the dependence between the joint probabilities (the entropy or the uncertainty) of two events. With respect to our task, the two events could be: what is the probability of an observer's response being a '*b*', given a stimulus letter '*h*'? Because there were 26 letters, the mutual information transmitted at a given letter slot ranged from zero bit for chance accuracy of 0.0384 to approximately 4.7 bits for perfect identification. To convert letter identification accuracy at a given letter slot to bits of mutual information transmitted, we used the following equation which was derived based on confusion matrices for single letter identification determined empirically by Beckmann (1998):

$$\text{Bits of information} = -0.037 + 4.676 \times \text{proportion correct of letter identification}$$

Then we summed up the total bits of information transmitted across all letter slots of the visual-span profile. This method of quantifying the visual span is akin to calculating the area under the curve.

2.5. Training task

The training task was similar to the one used in Chung (2007), whereby observers repeatedly identified the middle letters of trigrams rendered at a letter separation of 0.8× the standard spacing, which caused substantial crowding (letter identification accuracy was much lower than that for a larger letter separation). Letter size used for training was 1.4× the letter-size threshold. Regardless of the training group assignment, all observers completed six sessions of training. Each session consisted of 10 blocks of trials, with 100 trials per block.

2.6. Data analyses and reporting

Curve-fitting was accomplished using Igor Pro, which minimized the χ-square between the observed and predicted values. Unless otherwise stated, the reported error bars associated with the group-average values represent the 95% confidence intervals.

3. Results

Consistent with the results of a previous study (Chung, 2007), performance for identifying the middle letters of trigrams rendered at a letter separation of 0.8× the standard spacing at 10° in the normal periphery improved following 6 days of training for the daily training group. More importantly, using the same training task, observers in the weekly and biweekly training groups also showed substantial improvement following six sessions of training. Fig. 3 presents the data during training for two observers of each training group, one with the most (top panels) and the other with the least (bottom panels) amount of improvement, to show the range of performance during training for each training group. To quantify the improvement, for each observer, we fit a linear regression function relating his/her identification accuracy as a function of training block (Chung, 2007; Chung, Li, & Levi, 2012). 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. The *p*-values of this analysis are given in Table 1. After Bonferroni correction for multiple comparisons (p-value $\leqslant 0.0021$ to be considered as significant), the changes in performance during training for one observer of each of the three training groups are found to be not significantly different from 0, implying these three observers did not show any training effect. The proportion of observers who did not show any improvement following training is lower than those reported in the literature (Chung, Levi, & Tjan, 2005; Fahle & Henke-Fahle, 1996).

With the linear regression function, another analysis we performed to quantify the magnitude of improvement due to training was to compare the performance accuracy between the first and the last (60th) block based on the calculated values from the fitted function (see Table 1). Averaged across observers of each group, the magnitude of improvement ((accuracy for the last block − accuracy for the first block)/accuracy for the first block) was 37.9 ± 16.8%, 49.4 ± 24.5% and 26.1 ± 6.8% for the daily, weekly and biweekly training groups, respectively. These values are not different from each other (ANOVA: $F_{df=2} = 2.09$, $p = 0.15$; see Fig. 4A).

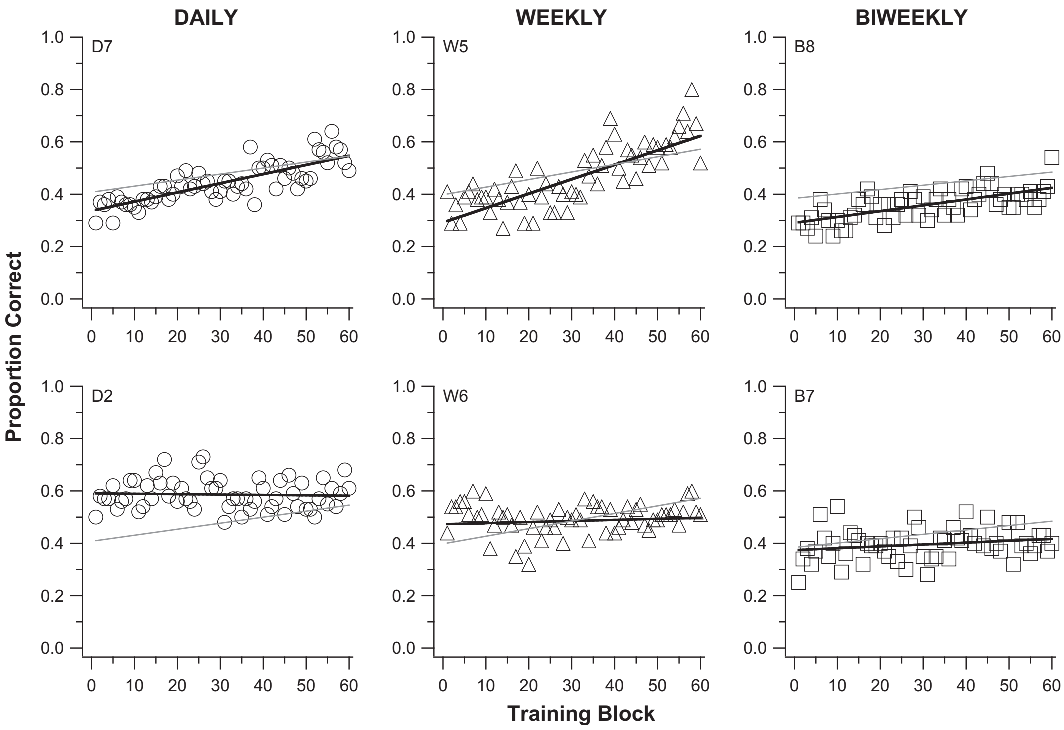

Fig. 3. Proportion of correct responses of identifying a letter flanked by two other letters at 0.8× the standard spacing is plotted as a function of training blocks, for two observers in each training group (left: daily training group; middle: weekly training group; right: biweekly training group). Each symbol represents the performance for a block of 100 trials. Data for the observer in each group demonstrating the most amount of improvement during training are shown in the top panels. Data for the observer in each group showing the least amount of improvement during training are shown in the bottom panels. In each panel, the black solid line represents the best-fit regression line to the 60 blocks of training data of the individual observer (observer's code given in the upper left corner); while the gray solid line represents the best-fit regression line to the data averaged across the eight observers in the respective training group.

Note that the initial performance was also similar across the three training groups (results not shown, ANOVA: $F_{df=2} = 0.13$, $p = 0.88$).

To examine whether the improvement in identifying crowded letters generalize to improved ability to identify letters flanked at larger letter separations, we compared the pre- and post-test measurements of the spatial extent of crowding. Fig. 5 shows the data from two observers of each training group, one with the largest change in the spatial extent of crowding (top panels) and the other with the smallest change in the spatial extent of crowding (bottom panels). Dotted lines in the panel for observer D3 represent how we defined the spatial extent of crowding (the letter separation that yielded a proportion-correct of identifying the middle flanked letters at 52% correct). The change in the spatial extent of crowding, expressed as the post-pre ratio (PPR: Levi & Li, 2009), averaged across observers of each group, is plotted for the three training groups in Fig. 4B (see also Table 1). A PPR smaller than 1 implies that the spatial extent is smaller after training than before. Averaged across observers of each group, the PPR averaged 0.815 ± 0.089, 0.806 ± 0.082 and 0.918 ± 0.063 for the daily, weekly and biweekly training groups, respectively. Because the 95% confidence ranges do not include the value of 1 (no improvement), these PPRs represent significant changes in the spatial extent of crowding following training for all three training groups. However, these changes are not different from one another (ANOVA: $F_{df=2} = 2.59$, $p = 0.10$).

We also examined if the improvement in identifying crowded letters generalize to an increased in the size of the visual span. Fig. 6 presents the data from two observers of each training group, one with the largest increase in the size of the visual span (top panels) and the other with the least increase in the size of the visual span (bottom panels). For each observer, we calculated the difference in the size of the visual span before and after training. Averaged across observers of each group, the visual span increased by 6.93 ± 2.22 bits, 7.37 ± 1.94 bits and 6.67 ± 1.85 bits for the daily, weekly and biweekly training groups, respectively (Fig. 4C and Table 1). In other words, all three groups exhibited significant improvement in the size of the visual span following training on identifying crowded letters (the 95% confidence limits of all the distributions did not contain the value of 0), but these improvements are not different among the three groups (ANOVA: $F_{df=2} = 0.12$, $p = 0.89$).

The third comparison we made for measurements before and after training was the size threshold for letter identification. The PPR for the three training groups averaged 0.78 ± 0.11, 0.81 ± 0.09 and 0.77 ± 0.06 for the daily, weekly and biweekly training groups, respectively (Fig. 4D and Table 1). ANOVA shows that these values are not statistically different from one another ($F_{df=2} = 0.23$, $p = 0.80$).

4. Discussion

Following six sessions (6000 trials) of repeated practice 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 letter separation at which the target letter was identified at 50% accuracy (after correction for guessing), decreased with training. Further, the improvement also led to an enlargement of the visual-span profile, and improved letter-size threshold. Most importantly, the amount of improvements for the trained task or other untrained tasks did not depend on whether training was conducted on a daily, weekly or biweekly schedule. These findings bear significant practical implication as they imply that there is no need for observers to attend daily

Table 1
p-Value of the regression fit to the training data (to determine if the slope differs from 0) and performance before and after training, are given for each observer in each training group. Error bars associate with the average values are ±95% CI.

Observer	Training data: *p*-value of regression fit	Training data: (60th block accuracy − 1st block accuracy)/1st block accuracy	Post/pre ratio of spatial extent of crowding	Post–pre difference in size of visual span (bits)	Post/pre ratio of letter-size threshold
Daily training group					
D1	0.0001	0.209	0.825	8.470	0.787
D2	0.7356	−0.015	0.948	−0.320	0.776
D3	<0.0001	0.635	0.592	5.469	0.970
D4	0.0004	0.308	0.700	7.008	0.990
D5	<0.0001	0.443	0.783	8.670	0.492
D6	<0.0001	0.639	0.982	7.491	0.679
D7	<0.0001	0.608	0.804	9.344	0.729
D8	0.0094	0.202	0.888	9.276	0.786
Average		0.379 ± 0.168	0.815 ± 0.089	6.926 ± 2.219	0.776 ± 0.110
Weekly training group					
W1	<0.0001	0.247	1.014	2.802	0.771
W2	<0.0001	0.908	0.789	4.920	0.832
W3	<0.0001	0.312	0.875	10.680	1.089
W4	<0.0001	0.571	0.774	7.697	0.688
W5	<0.0001	1.096	0.615	10.600	0.815
W6	0.3463	0.051	0.800	5.382	0.787
W7	<0.0001	0.277	0.869	8.965	0.685
W8	<0.0001	0.488	0.713	7.875	0.813
Average		0.494 ± 0.245	0.806 ± 0.082	7.365 ± 1.944	0.810 ± 0.087
Biweekly training group					
B1	0.0020	0.208	0.832	5.902	0.706
B2	0.0013	0.261	1.022	3.055	0.817
B3	0.0017	0.244	0.929	12.010	0.741
B4	0.0019	0.212	0.890	7.326	0.942
B5	<0.0001	0.327	0.853	6.022	0.682
B6	0.0008	0.273	0.925	6.380	0.716
B7	0.1078	0.112	1.075	4.526	0.831
B8	<0.0001	0.449	0.819	8.117	0.727
Average		0.261 ± 0.068	0.918 ± 0.063	6.668 ± 1.849	0.770 ± 0.060

training sessions in the laboratory to maximize their benefits from perceptual learning. This is especially important if perceptual learning is going to be adopted as a treatment for clinical populations.

It is well known that the magnitude of improvement following perceptual learning depends on many aspects of the training regime, including the training task (Fine & Jacobs, 2002), the total number of practice trials (e.g. Tsodyks & Gilbert, 2004), the amount of practice within each training session (Aberg, Tartaglia, & Herzog, 2009; Hussain, Sekuler, & Bennett, 2009; Kumar & Glaser, 1993) and the distribution of practice across time (Aberg, Tartaglia, & Herzog, 2009; Molloy et al., 2012; Taub & Goldberg, 1973; Xue et al., 2011). Task specificity is a notable characteristic of perceptual learning. Even for tasks that are seemingly related, for example, letter identification and reading, the transfer of learning from a trained to an untrained task is usually not complete. Maximal improvement is always obtained using a training task that targets directly at the intended visual function (Yu et al., 2010). The choice of a training task should therefore, be based on the specific visual function for which an improvement is desired. Even with the most relevant task, how much training is necessary? For sensory visual tasks, a recent report suggests that as few as one trial per condition on the first day of training is sufficient to produce an improvement greater than that for the control group (Hussain, Sekuler, & Bennett, 2009), although maximal learning requires more practice trials, from a total of several hundreds (Fiorentini & Berardi, 1981; Karni & Sagi, 1993; Poggio, Fahle, & Edelman, 1992; Wright & Sabin, 2007) to even thousands (Li, Levi, & Klein, 2004), depending on the specific training task. Note that however, performance for some tasks may continue to improve with additional practice up to 40–50 h (Li, Klein, & Levi, 2009; Li, Provost, & Levi, 2007). But how much practice within each training session is needed? Aberg, Tartaglia, and Herzog (2009) report that for a Chevron discrimination task, a minimum of 400 trials per training session is required to produce improvements. However, numerous reports show that the distribution of the practice trials, instead of the total number of practice trials, is a more important factor governing the amount of learning. The distribution of practice trials has often been discussed in the context of *massed* vs. *spaced* practice. Massed practice refers to the condition when successive trials are delivered one after another in a continuous manner whereas spaced training refers to the condition when there is a short time interval between successive trials, even if the interval is in the order of a few seconds (Ramsay, Utrecht, & Alkema, 1967; Taub & Goldberg, 1973; for reviews, see Donovan & Radosevich, 1999; or Cepeda et al., 2006). Most studies report that spaced practice is often accompanied by a greater magnitude of improvement and a higher retention of learning than massed practice (for reviews, see Donovan & Radosevich, 1999; or Cepeda et al., 2006), a finding that has been attributed to the need for consolidation of learned information, memory enhancement and/or the reduction of neural repetition suppression (e.g. Karni & Sagi, 1993; Xue et al., 2011). However, at least in the visual perceptual learning literature, one aspect of the training regime that receives little, if any, attention is the frequency of training sessions. The frequency of training sessions differs from the massed vs. spaced training in that the former one refers to the time interval between successive training sessions while the latter one refers to the time interval between successive trials. If perceptual learning is to be adopted as a treatment, or a mean to improve visual functions for the clinical populations, it is important to know what is the maximum time interval between training sessions such that the maximal learning effect could still be observed. Using an uncrowded letter identification training task, we found that the improvement in observers' performance

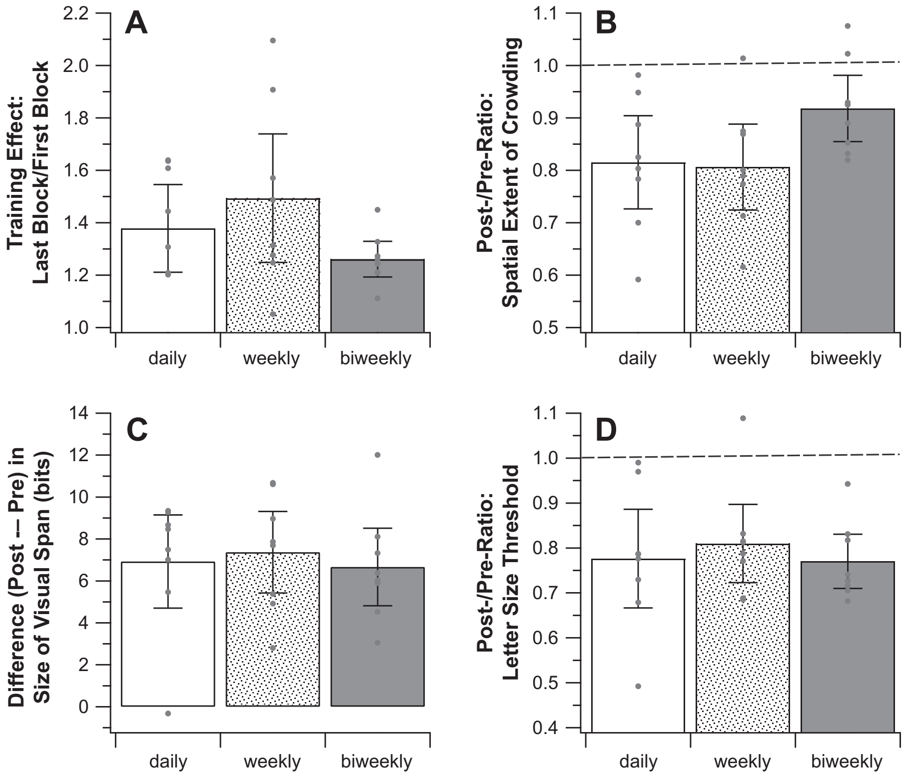

Fig. 4. Improvements in accuracy of identifying crowded letters during training (A) and the associated improvements in the spatial extent of crowding (B), the size of the visual-span profile (C) and letter-size threshold (D) are compared for the three training groups (daily, weekly and biweekly). Values plotted are averages across observers in each group. Error bars represent ±95% confidence intervals. Small gray dots represent values for individual observers. For (A), values plotted represent the ratio in performance accuracies between the last and the first block of trials. For (B) spatial extent of crowding and (D) letter-size threshold, post-/pre-ratios smaller than 1 (dashed lines) represent improvements.

for identifying crowded letters, and the improvements that were transferred to the untrained tasks (spatial extent of crowding, size of visual span and letter-size threshold) did not depend on whether observers were trained on a daily, weekly or biweekly schedule. In relation to the current theories of how perceptual learning relates to the need for the consolidation of learned information and memory, our results imply that the consolidation of the learned information is the same, and that the decay of the memory of the learned information does not change with the time interval between training sessions (at least up to intervals of 2 weeks). This seems to contradict our everyday experience that our memory of an event decays with time. However, our results are consistent with one or more of the following explanations: (1) the decay of memory of an event (can be considered as a "one-trial" practice) is different from the decay of memory of learned information through extensive practice (in our experiment, observers practiced 1000 trials per session), possibly due to a building-up or reinforcing of the representation of the learned information in memory through repeated practice in a single session; (2) the maximum amount of decay of learning occurs within 24 h after the end of a training session; and/or (3) the topping-up of the learning effect from each subsequent training session is the same regardless of the time interval since the last training session. Whether or not these speculations are correct is outside the scope of the current study but warrants further investigations.

Perceptual learning is defined as "...any relatively permanent and consistent change in the perception of a stimulus array, after practice or experience with this array" (Gibson, 1963). In order for perceptual learning to be a useful treatment to ameliorate the functional vision of clinical patients, in addition to the effectiveness of the training paradigm, another important consideration is the retention of the learning effect. Hence, an interesting question is whether the different training schedules affect how well observers retain their learning after training ceases, despite similar learning effects. Unfortunately, we did not examine retention in this study. However, in real life, if the trained task is an important daily task for the clinical patients (e.g. reading), then the patients are likely to be exposed to that task even after training ceases in the laboratory. This additional exposure could act as top-up training, and the question of how long the learning effect can be retained may not be too crucial an issue.

4.1. Comparison with Chung (2007)

The secondary goal of this study was to investigate if the improvements following a training protocol to learn to identify crowded letters would lead to an increase in the size in the visual span. The rationale for the investigation was to bridge our knowledge of the link between the visual span and reading speed (Legge, 2007; Legge, Mansfield, & Chung, 2001; Legge et al., 2007), and the lack of an accompanied increase in reading speed following a very similar training paradigm (Chung, 2007). Using essentially the same training protocol (same number of observers, same trained retinal location and same number of daily training sessions and number of trials), the magnitude of improvement during training ((performance accuracy for the last block – performance accuracy for the first block)/performance accuracy for the first block) was highly similar between the study of Chung (2007: 0.381 ± 0.086 [95% CI]) and the daily-training group of the present study (0.379 ± 0.168 [95% CI]).[3] Thus, we can assume that had we trained

[3] Note that highly similar amount of learning was obtained despite the different fonts, and how letter separations were specified in the current study vs. Chung (2007).

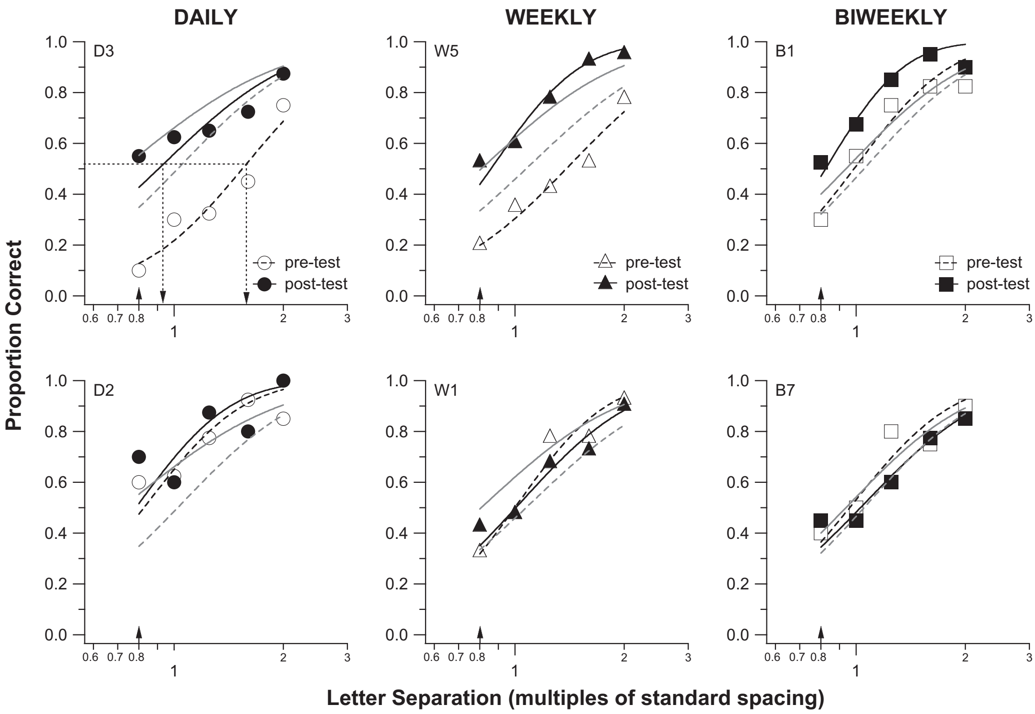

Fig. 5. Proportion of correct responses of identifying the middle letter of trigram is plotted as a function of letter separation (multiples of standard spacing) for two observers in each training group (left: daily training group; middle: weekly training group; right: biweekly training group), before (unfilled symbols) and after (filled symbols) training. Data for the observer in each group showing the most reduction in the spatial extent of crowding are shown in the top panels. Data for the observer in each group showing the least reduction in the spatial extent of crowding are shown in the bottom panels. Each set of data was fit with a cumulative-Gaussian function (smooth line through data points). 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, represented by the dotted lines shown in the upper left panel (observer D3). Arrows on the *x*-axes indicate the trained letter separation. In each panel, black lines represent the best-fit cumulative-Gaussian function to the individual observer's data; while gray lines represent the best-fit cumulative-Gaussian function to the data averaged across the eight observers in the respective training group.

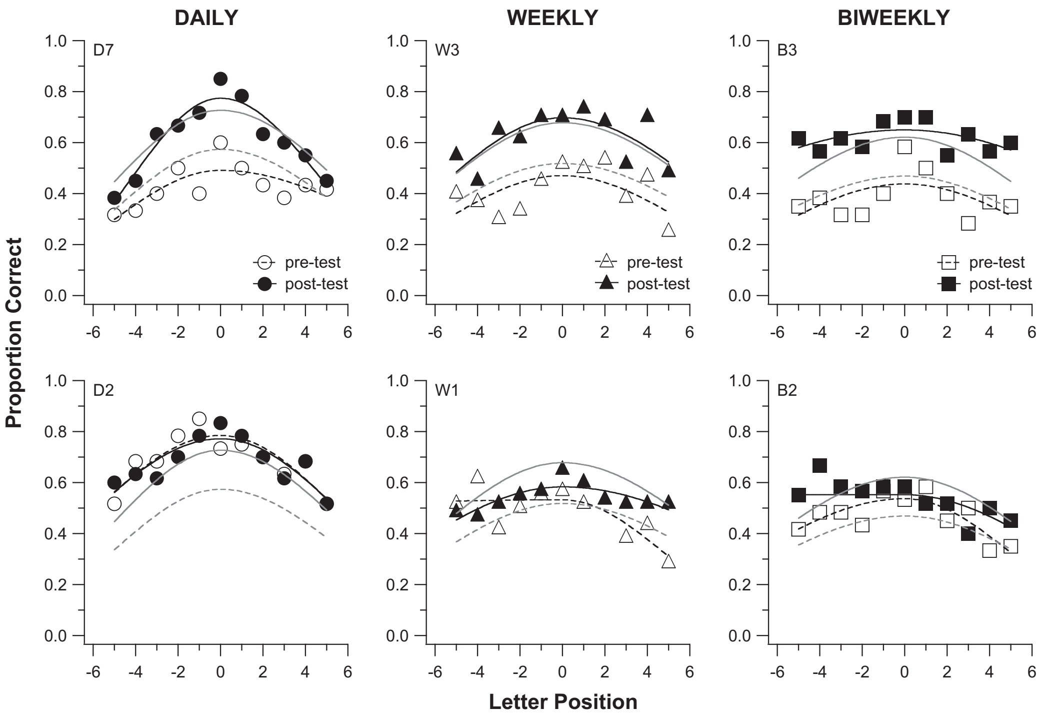

Fig. 6. Visual-span profiles, plots of proportion of correct responses of letter-recognition as a function of letter position, are shown for two observers in each training group (left: daily training group; middle: weekly training group; right: biweekly training group), before (unfilled symbols) and after (filled symbols) training. Data for the observer in each group showing the largest increase in the size of the visual span (see text for how the size of the visual span is quantified) are shown in the top panels. Data for the observer in each group showing the smallest increase in the size of the visual span are given in the bottom panels. Each set of data was fit with a split-Gaussian function (smooth line through data points). In each panel, black lines represent the best-fit split-Gaussian function to the individual observer's data; while gray lines represent the best-fit split-Gaussian function to the data averaged across the eight observers in the respective training group.

the same group of observers and measured their visual spans and reading speeds before and after training, they would likely show an enlargement of the visual span, as we found in the present study, but not an increase in the maximum reading speed, as was reported in Chung (2007). Previously, Legge and co-workers reported that the size of the visual span and the maximum reading speed exhibit the same qualitative dependence on character size, contrast (Legge et al., 2007) and testing eccentricity (Legge, Mansfield, & Chung, 2001). Further, following 20 repeated measurements of the visual-span profile, the size of visual span increases and is accompanied by an increase in the maximum reading speed (Chung, Legge, & Cheung, 2004). Based on the relationships between the size of the visual span and reading speed for different stimulus or testing conditions, Legge et al. (2007) deduced that an increase in the size of the visual span by 4.7 bits corresponds to approximately 39% increase in the maximum reading speed. However, despite the 7-bit increase in the size of the visual span exhibited by the daily-training group of the present study, maximum reading speed increased by a mere 7.2% in Chung (2007). What could have accounted for the lack of an increase in reading speed despite a sizeable increase in the size of the visual span?

The words we used for our RSVP reading task were presented left-justified on the computer display, and contained different number of letters. As such, the letters of the words extended to different letter positions right of fixation. We speculate that reading speed would benefit from an increase in the bits of information transmitted by the visual span if the increase in bits was distributed evenly across the different letter slots of the entire visual-span profile. However, our training task might have selectively improved the letter identification ability of our observers at only one letter position, the one corresponding to the target middle letter of trigrams, the position that was directly below fixation. Potentially our training task might have caused a location-specific improvement in letter identification, which might not have been of much benefit to reading. To determine if our training paradigm indeed led to an improvement that was specific to the location of the middle letters of trigrams (at 10° directly below fixation), we calculated the difference in bits of information transmitted by the visual span before and after training, only for letter position 0 which corresponded to the position of the middle letters of trigrams during training, and normalized this value to the overall difference in the size of the visual-span profile before and after training (from five letter slots left of the vertical midline to five letter slots right of the vertical midline). For comparison, this ratio was also calculated for the six observers who were trained at 10° in the inferior visual field using a visual-span letter-recognition training task, which presumably should have caused a more uniform improvement in letter-recognition across all letter positions (Chung, Legge, & Cheung, 2004). Averaged across observers of the respective group, these ratios are 0.11 ± 0.03 [95% CI] and 0.08 ± 0.02 for the present study and that of Chung, Legge, and Cheung (2004), respectively. These similar values suggest that although we only trained observers at one retinal location in the present study, the improvements transferred to other untrained letter positions thus leading to a general improvement in letter-identification across different letter positions. More importantly, these results suggest that our training trigrams presented directly below fixation did not cause a location-specific improvement in letter recognition, and thus could not explain why there was a lack of an increase in reading speed despite an increase in the size of the visual span.

Another possibility is that although both the visual span and reading speed are limited by low-level sensory factors such as letter contrast, letter size and retinal eccentricity, reading is likely to be subjected to additional higher-level limiting factors. These factors include contextual cues and the global shape information of combinations of letters (combinations of chunks of letters within a word or the overall word-shape). With our uncrowd training task, observers were trained to ignore the two flanking letters so as to improve their performance to identify the flanked letters. Thus, this training task would not have helped observers to improve their ability to recognize the global shape of letter-combinations, which could be important in reading (Pelli & Tillman, 2007). This is different from the visual-span letter-recognition training task in which observers had to identify all three letters of a trigram on each trial. One way to bridge the gap between the two types of training tasks is to train observers on trigrams that are always presented at one location, but require observers to report the identities for all three letters. An alternative way is to train observers using trigrams that are presented at different letter positions left and right of the vertical midline, but only require observers to report the middle letter of each trigram.

4.2. Summary

Consistent with the findings of Chung (2007), practicing identification of letters closely flanked by two other letters in peripheral vision improves observers' performance for identifying such letters. This improvement transfers to other larger letter separations such that the spatial extent of crowding is reduced following training, and is accompanied by an increase in the size of the visual span and a smaller letter-size threshold. Most importantly, the magnitudes of these improvements do not depend on whether observers were trained on a daily, weekly or biweekly schedule. Our results imply that clinical patients may benefit from perceptual learning even if they receive training once a week or every other week. Indeed, Chung (2011) trained a group of observers with central vision loss due to macular disorders once a week using a reading task, and found a sizeable (an average of 53%) increase in reading speed following six sessions (weeks) of training. This result provides support that clinical patients with reduced vision can benefit from perceptual learning on a weekly (perhaps even less frequent) training schedule.

Acknowledgments

This research was supported by NIH research Grant R01-EY012810 and the UC Berkeley Undergraduate Research Apprentice Program (URAP). We thank Amit Arunkumar and Deyue Yu for their assistance in data collection, and Jean-Baptiste Bernard and Girish Kumar for their helpful comments.

References

Aberg, K. C., Tartaglia, E. M., & Herzog, M. H. (2009). Perceptual learning with chevrons requires a minimal number of trials, transfers to untrained directions, but does not require sleep. *Vision Research, 49*, 2087–2094.

Astle, A. T., Webb, B. S., & McGraw, P. V. (2011). Can perceptual learning be used to treat amblyopia beyond the critical period of visual development? *Ophthalmic and Physiological Optics, 31*, 564–573.

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

Ball, K. K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research, 27*, 953–965.

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

Beckmann, P. J. (1998). *Preneural factors limiting letter identification in central and peripheral vision*. Unpublished doctoral thesis. University of Minnesota.

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

Cepeda, N. J., Pashier, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*, 354–380.

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

Chung, S. T. L. (2011). Improving reading speed for people with central vision loss through perceptual learning. *Investigative Ophthalmology and Visual Science, 52*, 1164–1170.

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. (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.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*, 795–805.

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

Fahle, M., & Henke-Fahle, S. (1996). Interobserver variance in perceptual performance and learning. *Investigative Ophthalmology and Visual Science, 37*, 869–877.

Fine, I., & Jacobs, R. A. (2002). Comparing perceptual learning across tasks: A review. *Journal of Vision, 2*, 190–203.

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.

Gibson, E. J. (1963). Perceptual learning. *Annual Review of Psychology, 14*, 29–56.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature, 402*, 176–178.

Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science, 18*, 88–94.

Huckauf, A., & Nazir, T. A. (2007). How odgcrnwi becomes crowding: Stimulus-specific learning reduces crowding. *Journal of Vision, 7*(2), 18.1–18.12.

Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2009). How much practice is needed to produce perceptual learning? *Vision Research, 49*, 2624–2634.

Hussain, Z., Webb, B. S., Astle, A. T., & McGraw, P. V. (2012). Perceptual learning reduces crowding in amblyopia and in the normal periphery. *Journal of Neuroscience, 32*, 474–480.

Huurneman, B., Boonstra, F. N., Cox, R. F. A., Cillessen, A. H. N., & van Rens, G. (2012). A systematic review on 'foveal crowding' in visually impaired children and perceptual learning as a method to reduce crowding. *BMC Ophthalmology, 12*, 27.

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

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature, 365*, 250–252.

Kumar, T., & Glaser, D. A. (1993). Initial performance, learning and observer variability for hyperacuity tasks. *Vision Research, 33*, 2287–2300.

Legge, G. E. (2007). *Psychophysics of reading in normal and low vision*. Mahwah, NJ and London: Lawrence Erlbaum Associates.

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), 9.1–9.15.

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*, 725–734.

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*, 635–654.

Levi, D. M., & Li, R. W. (2009). Perceptual learning as a potential treatment for amblyopia: A mini-review. *Vision Research, 49*, 2535–2549.

Li, R. W., Klein, S. A., & Levi, D. M. (2009). Prolonged perceptual learning of positional acuity in adult amblyopia: Perceptual template retuning dynamics. *Journal of Neuroscience, 28*, 14223–14229.

Li, R. W., Levi, D. M., & Klein, S. A. (2004). Perceptual learning improves efficiency by re-tuning the decision 'template' for position discrimination. *Nature Neuroscience, 7*, 178–183.

Li, R. W., Ngo, C., Nguyen, J., & Levi, D. M. (2011). Video-game play induces plasticity in the visual system of adults with amblyopia. *PLoS Biology, 9*(8), e1001135.

Li, R. W., Provost, A., & Levi, D. M. (2007). Perceptual learning results in substantial recovery of positional acuity and visual acuity in juvenile amblyopia. *Investigative Ophthalmology and Visual Science, 48*, 5046–5051.

Maniglia, M., Pavan, A., Cuturi, L. F., Campana, G., Sato, G., & Casco, C. (2011). Reducing crowding by weakening inhibitory lateral interactions in the periphery with perceptual learning. *PLoS ONE, 6*(10), e25568.

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

Molloy, K., Moore, D. R., Sohoglu, E., & Amitay, S. (2012). Less is more: Latent learning is maximized by shorter training sessions in auditory perceptual learning. *PLoS ONE, 7*(5), e36929.

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. (2007). Parts, wholes, and context in reading: A triple dissociation. *PLoS ONE, 2*(8), e680.

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

Polat, U. (2009). Making perceptual learning practical to improve visual functions. *Vision Research, 49*, 2566–2573.

Polat, U., Schor, C., Tong, J. L., Zomet, A., Lev, M., Yehezkel, O., et al. (2012). Training the brain to overcome the effect of aging on the human eye. *Scientific Reports, 2*, 278.

Ramsay, R. W., Utrecht, L. C., & Alkema, D. (1967). The effect of massed versus spaced practice on auditory threshold. *Acta Psychologica, 26*, 227–232.

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

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*, 379–423. 623–656.

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), 1–14 (article no. 16).

Taub, E., & Goldberg, I. A. (1973). Prism adaptation: Control of intermanual transfer by distribution of practice. *Science, 180*, 755–757.

Tsodyks, M., & Gilbert, C. (2004). Neural networks and perceptual learning. *Nature, 431*, 775–781.

Wright, B. A., & Sabin, A. T. (2007). Perceptual learning: How much daily training is enough? *Experimental Brain Research, 180*, 727–736.

Xue, G., Mei, L., Chen, C., Lu, Z.-L., Poldrack, R., & Dong, Q. (2011). Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. *Journal of Cognitive Neuroscience, 23*, 1624–1633.

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*, 36–45.