The Pattern of Learned Visual Improvements in Adult Amblyopia


PURPOSE. Although amblyopia is diagnosed in terms of a monocular letter acuity loss, individuals typically present with deficits on a wide range of spatial tasks. Many of these deficits can be collapsed along two basic visual dimensions (visual acuity and contrast sensitivity) that together account for most of the variability in performance of the amblyopic visual system. In this study, this space was exploited, to target the main deficits and fully characterize the pattern of learned visual improvements in adult amblyopic subjects.

METHODS. Twenty-six amblyopic subjects (mean age, 39 ± 12 years) were trained on one of four tasks, categorized as either visual acuity (letter or grating acuity) or contrast sensitivity (letter or grating contrast) tasks. Performance was measured on all tasks before and after training, to quantify learning along each dimension and generalization to the other dimension. Performance in 35 visually normal subjects (mean, age 24 ± 5 years) was used to establish normal variation in visual performance along each dimension, against which the learned improvements in amblyopic subjects was compared.

RESULTS. Training on the contrast sensitivity tasks produced substantial within-task learning and generalization to measures of visual acuity. The learned improvements in performance after training on the letter acuity task were also substantial, but did not generalize to contrast sensitivity.

CONCLUSIONS. Mapping the pattern of learning onto the known deficit space for amblyopia enabled the identification of tasks and stimulus configurations that optimized learning, guiding further development of learning-based interventions in this clinical group. (Invest Ophthalmol Vis Sci. 2011;52:7195–7204) DOI:10.1167/iovs.11-7584

Amblyopia is a common developmental visual disorder that is diagnosed by reduced visual acuity in one, or occasionally both, eyes—despite full optical correction and no evident ocular pathology. It affects approximately 2% to 4% of the population and remains the most common form of monocular vision loss in children, accounting for most children’s eye appointments in the United Kingdom. Orthodox treatment for amblyopia involves occlusion of the good eye, typically with a patch for lengthy periods. Amblyopia is caused by the disruption of normal visual input during the critical period(s) of visual development—postnatal windows of experience-dependent neural plasticity. The neural site of the amblyopic deficit is widely thought to be the primary visual cortex, although extrastriate areas may have a supplementary role. Until relatively recently, the adult visual cortex had not been considered capable of retaining any of the experience-dependent neural plasticity, which is so prominent during early visual development. However, it is becoming increasingly clear that altering visual experience, either by occluding the non-amblyopic eye or repeated exposure to controlled visual stimulation, can reshape brain function throughout the lifespan (see Ref. 8 for a review).

A much studied behavioral manifestation of neural plasticity in normal vision is perceptual learning, in which repeatedly practicing a challenging task can lead to substantial and long-lasting improvements in visual performance. Work on perceptual learning spans more than 150 years, and its effects have been widely documented in adulthood, well beyond the critical period(s) of development. In visually normal adults, perceptual learning has been shown to improve performance on a wide range of visual tasks, its hallmark being that improvements in performance are strongly coupled to trained visual attributes, such as the orientation of a stimulus. In contrast to the task-specific learning found in subjects with normal vision, trained improvements in amblyopic visual performance generalize to untrained tasks and novel stimuli, including visual acuity, visual counting, and stereoaucity. Generalization of perceptual learning to untrained tasks is key to harnessing this form of plasticity as an effective treatment for amblyopia, whether as a primary intervention or supplementary to traditional methods, such as occlusion therapy.

Although amblyopia is diagnosed in terms of a monocular letter acuity loss, individuals typically present with deficits on a wide range of spatial tasks. Many of these deficits can be collapsed along two basic visual dimensions that together account for virtually all the variation in performance of the amblyopic visual system. A factor analysis of visual function in 427 individuals with amblyopia, each of whom performed a range of spatial tasks, revealed two orthogonal dimensions of variation: visual acuity and contrast sensitivity. Subjects with different presumed etiologies occupied unique positions within this space, largely related to their degree of residual binocularity. In this study, we exploited this space (shown in Fig. 1), to target the main visual deficits and characterize the pattern of learned improvements and generalization in adult amblyopic subjects.

METHODS

Subjects

Thirty-five visually normal subjects and 26 subjects with naturally occurring amblyopia participated. The age of the normal and amblyopic subjects ranged from 16 to 38 (mean, 24 ± 5) and 17 to 63 (mean, 39 ± 12) years, respectively. All subjects who participated provided written informed consent. The experimental procedures adhered to...
the tenets of the Declaration of Helsinki and were approved by a local ethics committee at the School of Psychology, The University of Nottingham.

Subjects underwent a full ocular examination, refraction to determine their optical correction, and a full assessment of their binocular status, including a cover test. Best-corrected visual acuity was measured in each eye with the Bailey-Lovie chart.19 The subjects were classified as amblyopic if they had a visual acuity difference between the two eyes of at least 0.2 logMAR that could not be eliminated optically and had no evident ocular pathology. This is a widely accepted definition of amblyopia,20 which accommodates test-retest variability in visual acuity measurements.21 Twenty-one people did not meet our criteria for inclusion in the study. Table 1 shows clinical details for the amblyopic subjects.

Apparatus

Each subject’s head was held in position with a forehead and chin rest. Unless otherwise stated, all stimuli for the following psychophysical tests were generated on a computer with custom software written in Python22 and presented on a gamma-corrected CRT monitor (Vision Master Pro 514; Illya, Oude Meer, The Netherlands) with a refresh rate of 85 Hz and resolution of 1024 × 768 pixels. A digital-to-analog converter (Bits+; Cambridge Research Systems, Cambridge, UK) was used to increase the dynamic range of the luminance levels from 256 gray levels (standard 8-bit graphics card) to 16,384 levels (14-bit). A digital-to-analog converter (Bits+; Cambridge Research Systems, Cambridge, UK) was used to increase the dynamic range of the luminance levels from 256 gray levels (standard 8-bit graphics card) to 16,384 levels (14-bit). Unless otherwise stated, all screen-based tests were performed in a dark room, and the screen was viewed via a mirror at a distance of 6 m.

Stimuli

Letters. Landolt C’s were 5 × 5 sans serif letters. Gap width was equal to the stroke width and one fifth of letter width. Five Landolt C’s were arranged in a row (Fig. 2), each letter randomly oriented in one of the four cardinal directions, spaced half a letter width away from the next letter, and surrounded by a crowding bar to control for contour interactions.23 The width of the crowding bar was equal to the stroke width of the Landolt C’s and spaced a half letter width away from the letters. Uniform white (luminance, 180 cd m⁻²) and gray (luminance, 90 cd m⁻²) backgrounds were used for the letter acuity and letter contrast tasks, respectively.

Gabor. Gabor patches (example shown in Fig. 2) consisted of a horizontal sinusoidal carrier modulated on a uniform background (90 cd m⁻²), windowed by a two-dimensional Gaussian function (SD 0.5°). The mathematical expression describing a Gabor is:

\[ I(x, y) = I_m \left[ 1 + C_p \cos(2 \pi f_c x + \theta) \right] \times \exp \left[ -\frac{1}{2} \left( \frac{x}{\sigma_x} \right)^2 - \frac{1}{2} \left( \frac{y}{\sigma_y} \right)^2 \right] \]

where \( I_m \) is the mean luminance of the display, \( C_p \) is the peak contrast of the Gabor, \( f_c \) is the carrier spatial frequency, and \( \sigma_x \) and \( \sigma_y \) are standard deviations of the Gaussian envelope.

Tasks

Letter Acuity. Moving from left to right along a row of five Landolt C’s (Fig. 2), subjects indicated with a key press the orientation of the gap. The target letter was indicated by a marker above it (two letter widths above, two thirds letter height, one third letter width). On completing a line, the size of all letters was reduced in logarithmic steps, and letter acuity was scored in logMAR units on a letter-by-letter basis. Each letter scored 0.02 logMAR and a letter-by-letter (complete-line) termination criterion of four mistakes was used.24

Letter Contrast. Subjects viewed the screen at 1.5 m (except subject KT who was tested at 1 m). With the following exceptions, the stimulus configuration and judgment were the same as those used for the letter acuity task. Letter size was fixed at 1.30 logMAR (1.50 logMAR for KT), well above the acuity limit of all amblyopic subjects, ensuring that any deficit measured is likely to be the result of impaired contrast sensitivity rather than spatial resolution. The Michelson contrast (equation 2) was varied in logarithmic steps (Fig. 2). Performance on the task was scored in log 1/contrast units, such that a higher positive score corresponded to a lower contrast level and better performance (analogous to the Pelli-Robson chart scoring system).25 Each letter scored 0.02 log 1/contrast, and a letter-by-letter (complete-line) termination criterion of four mistakes was used.24

Grating Acuity. Horizontally oriented Gabor patches (envelope SD, 0.54°), were used to measure grating acuity (Fig. 2). In a temporal two-alternative, forced-choice task, subjects indicated which of two intervals contained the Gabor. On each trial, one randomly chosen interval contained the Gabor on a mean luminance background; the other interval contained the mean luminance background alone. A brief tone indicated the onset of an interval, each of which lasted 200 ms and were separated by 500-ms interstimulus intervals containing a fixation spot on a mean luminance background. Gabors were presented at 80% Michelson contrast (equation 2) for 180 ms in a raised cosine temporal window (10-ms onset and offset).

\[ \frac{(L_{\text{max}} - L_{\text{min}})}{L_{\text{max}} + L_{\text{min}}} \]  (2)

where \( L_{\text{max}} \) is the maximum luminance and \( L_{\text{min}} \) is the minimum luminance.

Spatial frequency was varied using a three-up, one-down staircase procedure that terminated after six reversals. Auditory feedback was provided for incorrect responses after each trial. Grating acuity thresholds were estimated as the geometric mean of the last four reversals.

Grating Contrast. With the following exceptions, the timing, procedure, and staircases were the same as those used for grating acuity. Gabor patches, were used to measure contrast sensitivity in a range of spatial frequencies (0.5, 1, 2, 4, 8, 16, and 32 cyc/deg). Contrast was varied in Michelson units (equation 2), and contrast sensitivity was calculated as the reciprocal of the geometric mean of the contrast threshold for the last four reversals. Staircases for different frequencies were randomly interleaved and terminated once all staircases had completed six reversals.

Training

Before and after training, we measured the subjects’ performance on all four of the tasks described above. Task order was randomized and the pre- and posttraining phases each lasted approximately 90 minutes. During training, subjects were randomly assigned to train on one of the tasks for 10 daily sessions. Five visually normal subjects trained on each task. Table 2 shows, for each test, how many amblyopic subjects trained, the approximate number of trials completed, the number of runs completed per training session, and the approximate duration of each training session. The number of trials per run and the duration...
of each training session were dependent on the performance and termination criteria of each test. Training sessions were separated by no more than 3 days.

Fifteen observers with normal vision did not undergo training and instead had performance measured on each task during a test session, followed by repeat measurement of performance during a retest session 2.5 weeks later. The visual acuity of all subjects was measured before and after training with the Bailey-Lovie chart, as well as the computer-based Landolt C letter acuity test.

All testing and training was performed with the ambyopic eye of amblyopic observers or the non-dominant eye in normal observers. All subjects were optically corrected for the working distances used.

### Data Analysis

Using 30 visually normal subjects (those who did not train on the task), we derived test and retest confidence intervals (CIs) in sessions separated by 2 weeks for each of the four tasks. The 95% CIs were calculated by taking the 2.5% and 97.5% percentiles from 10,000

<table>
<thead>
<tr>
<th>Subject’s</th>
<th>Task</th>
<th>Age</th>
<th>M/F</th>
<th>Amb Eye</th>
<th>Refractive Error</th>
<th>VA</th>
<th>Strab</th>
<th>Amblyopia Type</th>
<th>Treatment History</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC</td>
<td>GA/LC</td>
<td>42</td>
<td>F</td>
<td>L</td>
<td>R pl/−0.75 × 85</td>
<td>0.00</td>
<td>LSOT</td>
<td>Strab and aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>NGM</td>
<td>GA</td>
<td>29</td>
<td>M</td>
<td>R</td>
<td>R +3.00/−1.00 × 60</td>
<td>0.32</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>PCP</td>
<td>GA</td>
<td>40</td>
<td>M</td>
<td>R</td>
<td>R +3.75/−0.75 × 55</td>
<td>0.24</td>
<td>RSOT</td>
<td>Strab and aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>RLC</td>
<td>GA</td>
<td>29</td>
<td>M</td>
<td>L</td>
<td>R +0.25 DS</td>
<td>0.00</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>MP</td>
<td>GA/LC</td>
<td>52</td>
<td>M</td>
<td>R</td>
<td>R +4.50/−1.00 × 80</td>
<td>0.70</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs</td>
</tr>
<tr>
<td>LE</td>
<td>GA/LC</td>
<td>54</td>
<td>M</td>
<td>L</td>
<td>R −5.25/−0.50 × 85</td>
<td>0.00</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>TS</td>
<td>GA</td>
<td>45</td>
<td>F</td>
<td>L</td>
<td>R pl</td>
<td>0.04</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>JW</td>
<td>GA</td>
<td>55</td>
<td>F</td>
<td>L</td>
<td>R +0.75/−0.25 × 75</td>
<td>0.00</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs</td>
</tr>
<tr>
<td>CM</td>
<td>LA</td>
<td>24</td>
<td>F</td>
<td>L</td>
<td>R +0.25 DS</td>
<td>−0.08</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>WT</td>
<td>LA</td>
<td>44</td>
<td>F</td>
<td>L</td>
<td>R +0.25 DS</td>
<td>0.00</td>
<td>LSOT</td>
<td>Strab and aniso</td>
<td>Strab op, specs and occlusion</td>
</tr>
<tr>
<td>CH</td>
<td>LA</td>
<td>50</td>
<td>M</td>
<td>L</td>
<td>R +3.00/−2.00 × 10</td>
<td>1.30</td>
<td>L micro</td>
<td>Strab</td>
<td>None</td>
</tr>
<tr>
<td>AG</td>
<td>LA</td>
<td>42</td>
<td>M</td>
<td>R</td>
<td>R −4.00/−1.00 × 165</td>
<td>0.50</td>
<td>RXOT</td>
<td>Strab</td>
<td>None</td>
</tr>
<tr>
<td>ACT</td>
<td>LA</td>
<td>31</td>
<td>F</td>
<td>L</td>
<td>R +3.00/−1.50 × 45</td>
<td>0.06</td>
<td>LSOT</td>
<td>Strab and aniso</td>
<td>Strab and occlusion</td>
</tr>
<tr>
<td>AS</td>
<td>GC</td>
<td>47</td>
<td>M</td>
<td>L</td>
<td>R −1.75/−1.00 × 30</td>
<td>0.10</td>
<td>LSOT</td>
<td>Strab and aniso</td>
<td>Strab and occlusion</td>
</tr>
<tr>
<td>KAN</td>
<td>GC</td>
<td>63</td>
<td>M</td>
<td>L</td>
<td>R +1.50/−0.25 × 130</td>
<td>0.80</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>SB</td>
<td>GC</td>
<td>37</td>
<td>F</td>
<td>L</td>
<td>R −1.00 DS</td>
<td>0.00</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>EF</td>
<td>GC</td>
<td>45</td>
<td>M</td>
<td>R</td>
<td>R +5.50/−2.00 × 130</td>
<td>0.60</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>CMW</td>
<td>GC</td>
<td>22</td>
<td>M</td>
<td>R</td>
<td>R +4.75 DS</td>
<td>0.60</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>FH</td>
<td>GC</td>
<td>33</td>
<td>F</td>
<td>R</td>
<td>R +3.00/−2.00 × 135</td>
<td>0.32</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs</td>
</tr>
<tr>
<td>BM</td>
<td>LC</td>
<td>27</td>
<td>F</td>
<td>L</td>
<td>R +3.75/−0.50 × 140</td>
<td>0.24</td>
<td>LSOT</td>
<td>Strab and aniso</td>
<td>Strab and occlusion</td>
</tr>
<tr>
<td>CB</td>
<td>LC</td>
<td>21</td>
<td>F</td>
<td>R</td>
<td>R +4.00/−0.25 × 110</td>
<td>0.14</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>JN</td>
<td>LC</td>
<td>17</td>
<td>F</td>
<td>R</td>
<td>R +3.00/−1.00 × 100</td>
<td>0.20</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
<tr>
<td>KR</td>
<td>LC</td>
<td>56</td>
<td>M</td>
<td>L</td>
<td>R +2.00/−0.50 × 10</td>
<td>−0.24</td>
<td>LXOT</td>
<td>Strab and aniso</td>
<td>Strab op, specs and occlusion</td>
</tr>
<tr>
<td>KT</td>
<td>LC</td>
<td>26</td>
<td>F</td>
<td>R</td>
<td>R −4.00/−2.00 × 10</td>
<td>1.35</td>
<td>R Microt</td>
<td>Strab and aniso</td>
<td>Strab and occlusion</td>
</tr>
<tr>
<td>RL</td>
<td>GC/LC</td>
<td>47</td>
<td>M</td>
<td>R</td>
<td>R +8.00/−2.50 × 18</td>
<td>0.26</td>
<td>NMD</td>
<td>Bilateral hypermetropic Aniso</td>
<td>Specs</td>
</tr>
<tr>
<td>IMB</td>
<td>LC</td>
<td>41</td>
<td>M</td>
<td>L</td>
<td>R +0.50 DS</td>
<td>0.10</td>
<td>NMD</td>
<td>Aniso</td>
<td>Specs and occlusion</td>
</tr>
</tbody>
</table>

M, male; F, female; Amb, amblyopia; Strab, strabismus; GA, grating acuity; LA, letter acuity; GC, grating contrast; LC, letter contrast; R, right; L, left; pl, plano; VA, initial visual acuity (logMAR with Bailey-Lovie); NMD, no movement detected; RSOT/LSOT, right/left esotropia; RXOT/LXOT, right/left exotropia; Aniso, anisometropia; Specs, spectacles; R/L tropia, right hypermetropia; L/R, left hypermetropia; microt, microtropia. Subjects DC, MP and LE were trained on GA then LC.

Subject RL (bilateral amblyopia) RE was trained.
bootstrapped samples (using the bias-corrected accelerated-percentile method) of the ratio between test and retest scores for these subjects.

We expressed changes in performance on a task across different training sessions as the ratio between performance on each training day and the first day of training. We compared the trained within- and between-task performance improvements to the untrained test and retest bootstrapped CIs for that task (described above). To determine whether within- and between-task improvements in performance were significantly different between trained amblyopic groups and non-trained normal subjects, we performed Monte Carlo permutation testing (using 10,000 resamples).

To characterize performance on the grating contrast task, the area under the log contrast sensitivity function (AULCSF) was calculated by fitting a third-order polynomial to the log contrast sensitivity versus log spatial frequency data of each subject and integrating between the lowest and highest spatial frequencies. A similar method has been used to compare changes in contrast sensitivity.26

**RESULTS**

We first considered the variability of taking repeated performance measurements on a task and then compared the estimates to training-induced changes in performance on the same task. This method enabled us to infer whether training had produced perceptual improvements greater than the inherent variability in repeated measurements of performance. Figure 3 shows the mean scores and 95% CIs for test and retest measures on each task. Performance was better (on average, by 3%) and less variable on most tasks at retest, possibly because of procedural effects.

Figure 4 shows example learning curves for normal and amblyopic subjects on each of the four trained tasks. Figures 4A and 4B show letter-based tasks, and Figures 4C and 4D show grating-based tasks. Mean performance and CIs of normal subjects who did not train on the tasks are shown by dashed lines.

**TABLE 2. Number of Amblyopic Subjects Trained on and Approximate Number of Trials Completed for Each Task, along with the Number of Runs Completed and Duration of Each Training Session**

<table>
<thead>
<tr>
<th>Trained Task</th>
<th>Amblyopic Subjects Trained (n)</th>
<th>Trials (−n)</th>
<th>Runs of Test per Training Session</th>
<th>Approximate Time per Training Session (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letter acuity</td>
<td>5</td>
<td>2000</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Letter contrast</td>
<td>10</td>
<td>3795</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Grating acuity</td>
<td>8</td>
<td>3300</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Grating contrast</td>
<td>6</td>
<td>4200</td>
<td>1</td>
<td>30</td>
</tr>
</tbody>
</table>

The number of trials per run and the duration of each training session were dependent on the performance and termination criteria of each test.
horizontal lines and gray shaded regions at sessions 1 and 12. Normal subjects (filled circles) showed modest improvements on all the tasks over the course of training. In contrast, amblyopic subjects (open circles) started with poor performance when compared with their normal counterparts, improved significantly during the course of training, and in some instances reached performance levels that were within the normal range. For example, a normal subject (DPM) showed little change in performance over the course of letter acuity training, with performance remaining at the normal level of the task throughout, whereas an amblyopic subject (ACT) had relatively poor visual acuity at the start of training, but improved by more than 3 lines (0.32 logMAR) over the course of training. In fact, by the final session, their acuity was better than 0.00
logMAR, equivalent to the mean performance of visually normal subjects (−0.05 logMAR). On the grating acuity task, however, we found little or no learning in both the normal and the amblyopic subjects.

We expressed learned improvements in performance relative to performance before training (post-/pretraining ratio for the letter tests and pre-/posttraining ratio for the grating tests, hereafter referred to as PPR), where numbers less than one constitute learning. Group mean PPR scores (±SEM) were calculated for each of the tasks and are shown in Figure 5. Since the performance on the letter acuity task is expressed in logMAR units, some scores are negative. This is problematic when calculating ratios such as the PPR. To circumvent the problem, we converted letter acuity scores to MAR and letter contrast scores converted into raw Michelson contrast units before PPR was calculated.

Normal subjects showed limited improvements in performance over the course of training: Those who trained on the letter contrast task improved by the largest amount (mean PPR, 0.65; SEM ± 0.07), followed by letter acuity (0.82 ± 0.08), grating contrast (0.90 ± 0.05), and grating acuity (0.93 ± 0.06). The improvements found for normal subjects trained on these tasks were not significantly different from the changes in performance found for normal subjects who did not train, apart from those trained on the letter contrast task (P = 0.0011). Amblyopic subjects, on the other hand, improved more than normal subjects in all the tasks apart from grating acuity, where both groups showed little or no change in performance. Amblyopic subjects who trained on letter contrast improved the most (PPR, 0.41 ± 0.09), followed by letter acuity (PPR, 0.66, ± 0.06), and grating contrast (PPR, 0.75 ± 0.04) over the period of training. Apart from the grating acuity task, the improvement of all amblyopic groups significantly exceeded the change in performance of normal subjects who did not train (all P < 0.001).

Figure 6 shows individual amblyopic data. Posttraining performance is plotted against performance before training for each of the tests. Points lying outside of the shaded region correspond to subjects whose performance improved after training. Taken together, these data suggest that contrast sensitivity is much more amenable to learning in both normal and amblyopic subjects than are acuity-based tasks.

Even though subjects were randomly assigned to the training groups, we wondered whether the composition of these groups may have contributed to the differential levels of learning on each task. Since neural plasticity is thought to dissipate over the course of the lifespan, the age of subjects could be confounded by the amount of learning. We therefore compared the age of the subjects in each group (Table 3). Amblyopic subjects trained on the letter contrast task had the lowest mean age of the amblyopic groups and showed the greatest amount of learning on the trained task. However, there was no statistically significant correlation between age and the magnitude of improvement on the trained task for these subjects (r<sub>27</sub> = 0.26; P = 0.17). Alternatively, the different levels of learning between the groups might be related to their respective visual acuities before training. Even though mean acuity before training was poorest in the amblyopic groups who improved most (contrast trained), again there was no correlation between Bailey-Lovie visual acuity (which no subjects trained on) and the magnitude of improvement in these subjects (r<sub>27</sub> = −0.18; P = 0.34).

**Figure 5.** Normalized mean learning curves for each of the tasks. Mean normalized performance is shown for each of the tasks for subjects with normal (●) and amblyopic (○) vision. Learned improvements in performance are expressed relative to performance on the first day of training. Points lying below the solid horizontal line (PPR = 1) indicate an improvement beyond the level of performance on the first day of training. Dashed horizontal lines: mean test–retest ratio of normal subjects who did not undergo training on the task (gray shaded region shows 95% CI). Error bars, SEM.
In the amblyopic subjects trained on the letter contrast task, the level of improvement found correlated with the initial performance on the task \((r_p = 0.76; P < 0.05)\). This relationship is shown in Figure 7. Greater levels of improvement are found in subjects with poorer starting performance compared with those with better starting performance.

Thus far, only improvements in performance on the trained task have been considered (within-task learning). We now considered how these improvements generalize to untrained tasks (between-task learning). Figure 8 shows the transfer of learning for both normal and amblyopic subjects to all untrained tasks. We considered only transfer between tasks that converge on similar performance levels.

Each shows the average trained improvement on each task (bars in lower contrast) and how these transferred to the respective letter or grating task on the other dimension. Subjects with normal vision showed modest amounts of transfer to other tasks. Amblyopic subjects who trained on letter contrast, not only improved significantly on the task itself, but also improved on letter acuity (beyond the change in performance found in normal observers who did not undergo training).

This pattern of transfer is represented in Figure 9, where improvements on each task are plotted in the acuity-contrast space. Changes in visual acuity are shown on the abscissas and changes in contrast sensitivity on the ordinates. Shaded regions

![Figure 6](image)

**Figure 6.** Performance for individual amblyopic subjects as a function of performance before training. Subjects with results lying on the diagonal showed no benefits from training. Axes have been oriented such that points lying outside the shaded region denote subjects whose performance improved on the task after training. (B) Gray filled circles correspond to subjects previously trained on the grating task.

![Figure 7](image)

**Figure 7.** Improvement in letter contrast plotted against start letter contrast, for amblyopic subjects trained on the letter contrast task with linear regression curve fit. The two variables correlated strongly \((r_p = 0.76; P < 0.05)\).

<table>
<thead>
<tr>
<th>Group/Trained Task</th>
<th>Mean Age (±SD)</th>
<th>Mean Visual Acuity before Training (±SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter acuity</td>
<td>24 (5)</td>
<td>−0.08 (0.09)</td>
</tr>
<tr>
<td>Letter contrast</td>
<td>26 (5)</td>
<td>−0.04 (0.08)</td>
</tr>
<tr>
<td>Grating acuity</td>
<td>25 (4)</td>
<td>−0.08 (0.07)</td>
</tr>
<tr>
<td>Grating contrast</td>
<td>25 (5)</td>
<td>−0.04 (0.10)</td>
</tr>
<tr>
<td>Amblyopic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter acuity</td>
<td>38 (10)</td>
<td>0.50 (0.50)</td>
</tr>
<tr>
<td>Letter contrast</td>
<td>38 (14)</td>
<td>0.64 (0.38)</td>
</tr>
<tr>
<td>Grating acuity</td>
<td>43 (10)</td>
<td>0.50 (0.28)</td>
</tr>
<tr>
<td>Grating contrast</td>
<td>42 (14)</td>
<td>0.67 (0.21)</td>
</tr>
</tbody>
</table>

![Figure 8](image)

**Figure 8.** Bar charts showing improvements in performance on the trained task and transfer of learning to the untrained task. Each bar graph shows data for subjects who trained on one of the tasks. For example, (A) shows data for subjects who trained on the letter acuity task. It shows improvements on the trained task (letter acuity) and transfer of improvements in performance to the letter contrast task for these subjects. Results corresponding to within-task learning are displayed in lower contrast.
represent 95% CI for the respective dimensions. The origin in Figure 9 represents no learning along both dimensions (PPR = 1) and the initial performance level of all amblyopic groups before training. The axes are oriented such that, if a group of subjects improves in visual acuity over the course of training, the data point representing the group will move rightward in this space; improvements in contrast sensitivity, on the other hand, will produce upward movement.

The most notable features of the amblyopic data are that training on contrast-based tasks confers significant visual benefits along both dimensions, whereas letter acuity training produces benefits that are tightly coupled to the trained dimension. We wondered whether the lack of improvement on the grating acuity task was due to a characteristic of the task itself or had something to do with the subjects who were assigned to train on the task. Therefore, we asked three amblyopic subjects who had trained on the grating acuity test (LE, MP, DC) to return and train on the letter contrast task. Even though these subjects showed no improvement on the grating acuity test, they showed learning on the letter contrast task and transfer of learning to the letter acuity test beyond the 95% CI for normal subjects who did not train on these tasks. Adding the data for these participants to the results of those previously trained on the letter contrast task had little effect on the magnitude of within-task learning (PPR = 0.39) or the degree of transfer to the letter acuity task (PPR = 0.80) and did not change the pattern of improvements found.

**DISCUSSION**

We defined the visual deficits of adult amblyopic subjects using two dimensions of variation: visual acuity and contrast sensitivity. Using this deficit space, we found that amblyopic subjects showed greater learning within and across these dimensions relative to normal control subjects, benefiting most from training on a letter-based contrast task. Other studies have shown that amblyopic subjects can learn on a range of visual tasks, including positional acuity, contrast detection, and letter recognition (see Ref. 28 for a review). Transfer of learning has also been demonstrated to untrained tasks such as letter acuity, visual counting, and stereoaucity. However, previous learning studies have used a wide range of visual tasks, psychophysical procedures, stimuli, subject numbers, and different forms of experimental control, making it difficult to establish patterns of learning in adult amblyopia. To faithfully characterize the rich pattern of learned improvements in adult amblyopia, it is critical not only to consider the broader spectrum of visual deficits, but also to consider the generalization of learning within a framework that imposes a priori constraints on their interpretation.

In this study, we adapted a framework developed by McKee et al., which classifies the visual loss in amblyopic subjects within an acuity–contrast space. We used this space to characterize learning-induced changes in amblyopic visual performance relative to normal variation along each dimension. We designed tests that used common stimuli (letters and gratings) and similar test procedures to limit the impact of procedural learning on our results, facilitating comparison between test results. Mean and variation of normative scores for each of the tests were in good accordance with previous reports in the literature.

Quite often, improvement in amblyopic performance is compared to that in the fellow eye. This may not be a valid comparison since performance of one eye can change after training of the other eye. Moreover, the fellow eye of amblyopic subjects is not considered by many to be completely normal. To avoid this problematic comparison, we compared performance of amblyopic subjects to visually normal controls and calculated CIs for each task. This sets a minimum baseline and level of variation against which any improvements in visual performance can be compared. This approach is an alternative to showing that the slope of a learning function is significantly different from 0 (flat line)—a situation that compares performance changes to a baseline with no associated variation.

If asked to define the quality of vision using a single number, most clinicians would opt for an estimate of letter acuity. Because of the familiarity and sensitivity of this measure, it is also used as a key indicator of treatment success in amblyopia. Amblyopic subjects that trained on a letter acuity task showed a 34% change in letter score, which equates to an improvement of 0.2 logMAR. The maximum improvement for an amblyopic subject was more than 0.3 logMAR (ACT). An improvement of this magnitude would require around 380 hours of patching in a child. In the present study, we were able to generate these changes in a fraction of the time (<10 hours). Apart from the rapidity of learning, more interesting in many ways is the generalization of learning: Training on letter contrast generated levels of improvement on letter acuity similar to those generated by training on the task itself.
In contrast to two previous reports, we found no improvement in grating acuity for either normal or amblyopic subjects. One study that did show learning on this task used only a single subject, who had lost the fellow eye. It is likely that the mechanisms of visual recovery in this subject will be very different from those of our cohort. The other study found improvements in amblyopic subjects who had no previous treatment, but little or no improvements in previously treated subjects. Most of our grating acuity group (6/8) had undergone patching during childhood. These differences in treatment history may partially explain the lack of learning in this group. It may not be the full explanation, however, since many subjects in the other groups that did show learning had previously been treated with occlusion therapy. Alternatively, the capacity for improvement may be limited for grating acuity tasks in adulthood or fundamental differences between stimuli (broadband versus narrowband) and task (detection versus discrimination) may be important. Consistent with this notion, when gratings are used in a spatial frequency discrimination task, learning can be substantial.

Other studies (also using contrast based tasks) have achieved PPR ≥ 0.34 (PPR found for letter contrast task), although subjects received more training. We suspect that the very rapid rate of learning and degree of generalization in our subjects in the other groups that did show learning had previously been treated with occlusion therapy. From a clinical perspective, any therapeutic intervention that confers the largest magnitude of within dimension learning can be substantial.

We restricted our comparisons to those between tasks of equal precision (i.e., those between the two letter tasks and those between the two grating tasks only), as differences in task precision can influence the amount of learning and transfer found.

Differences in the amount of transfer found for each task may be due to differences in the number of trials completed per training session for each task. Greater levels of transfer are to be expected for briefer periods of training. The other study found. Differences in the amount of transfer found for each task may be due to differences in the number of trials completed per training session for each task. Greater levels of transfer are to be expected for briefer periods of training. However, we showed greater levels of transfer for contrast tasks on which subjects received more training. Therefore, the greater levels of transfer found for the contrast tasks are unlikely to be accounted for by the differential amount of training received.

One could argue that improvements in performance on tests are due to factors other than genuine perceptual improvements in performance. For example, better control of accommodation may play a role in the improvements found. However, the amount of accommodation necessary to focus on the test targets would have been minimal due to the distances used and the fact that subjects were optically corrected for the working distances used.

From a clinical perspective, any therapeutic intervention should optimize the magnitude, timescale, and generalization of learned visual improvements in amblyopia. Mapping the pattern of learning onto the known deficit space for amblyopia enabled us to identify which task best met these conditions for a fixed training period. Letter-based contrast training confers the largest magnitude of within dimension learning and across dimension generalization over very short time scales. This makes crowded contrast-based letter tasks ideal candidates for further development of learning-based interventions in this clinical group. Recently, we have also shown that this approach puts in place the necessary neural precursors required to fully recover stereo acuity in adult amblyopic subjects.

References