Transfer and retention of oculomotor alignment rehabilitation training

Kerri Walter Northeastern University, Boston, MA, USA

Yesenia Taveras-Cruz Northeastern University, Boston, MA, USA

Peter Bex Northeastern University, Boston, MA, USA

Ocular alignment defects such as strabismus affect around 5% of people and are associated with binocular vision impairments. Current nonsurgical treatments are controversial and have high levels of recidivism. In this study, we developed a rehabilitation method for ocular alignment training and examined the rate of learning, transfer to untrained alignments, and retention over time. Ocular alignment was controlled with a real-time dichoptic feedback paradigm where a static fixation target and white gaze-contingent ring were presented to the dominant eye and a black gaze-contingent ring with no fixation target was presented to the nondominant eye. Observers were required to move their eyes to center the rings on the target, with real-time feedback provided by the size of the rings. Offsetting the ring of the nondominant temporal or nasal visual field required convergent or divergent ocular deviation, respectively, to center the ring on the fixation target. Learning was quantified as the time taken to achieve target deviation of 2° (easy, E) or 4° (hard, H) for convergence (CE, CH) or divergence (DE, DH) over 40 trials. Thirty-two normally sighted observers completed two training sequences separated by one week. Subjects were randomly assigned to a training sequence: CE-CH-DE, CH-CE-DE, DE-DH-CE, or DH-DE-CE. The results showed that training was retained over the course of approximately one week across all conditions. Training on an easy deviation angle transferred to untrained hard angles within convergence or divergence but not between these directions. We conclude that oculomotor alignment can be rapidly trained, retained, and transferred with a feedback-based dichoptic paradigm. Feedback-based oculomotor training may therefore provide a noninvasive method for the rehabilitation of ocular alignment defects.

Introduction

Strabismus and binocular vision

Strabismus refers to a condition in which binocular axis alignment is impaired. If this impairment is not corrected during development, it may lead to visual dominance being highly favored in a single eye and decreased visual acuity in the other eye, known as amblyopia (Epelbaum et al., 1993). The current estimated prevalence rate for strabismus is between 2% and 5% (Beauchamp et al., 2003; Friedmann et al., 1980; Graham, 1974; Rutstein et al., 2010; Stidwill, 1997). Strabismus most commonly occurs in the horizontal axis. An eye that deviates inward (nasally) is considered esotropic, while an eye that deviates outward (temporally) is considered exotropic (Gunton, Wasserman, & Debenedictis, 2015). Approximately half of the adult population of strabismus patients received diagnoses as a child. These patients did not undergo treatment, had unsuccessful treatment, or relapsed back into ocular misalignment after treatment (Beauchamp et al., 2003). The probability of having strabismus is unaffected by ethnicity and social class (Graham, 1974).

Binocular coordination orients the eyes to fixate on the same point in space. Owing to the horizontal separation of the eyes, this generates interocular disparity that varies with object depth across the visual field (Ritter, 1977). To identify corresponding retinal points and estimate object depth, the visual system limits the horizontal feature search distance in the retina by approximately 1° (Harrold & Grove, 2015). Patients with strabismus have eyes that do not align on the same point in space, and so the corresponding features presented in each eye can lie outside of this search range (Read, 2015). This prevents the two images from being fused and can result in stereoblindness and diplopia (double vision). To avoid diplopia, the deviated

Citation: Walter, K., Taveras-Cruz, Y., & Bex, P. (2020). Transfer and retention of oculomotor alignment rehabilitation training. *Journal of Vision*, *20*(8):9, 1–16, https://doi.org/10.1167/jov.20.8.9.

https://doi.org/10.1167/jov.20.8.9 Received April 9, 2020; published August 6, 2020 ISSN 1534-7362

eye can be suppressed (Pratt-Johnson & Tillson, 1984), which prevents stereopsis.

Quality of life

When there is no diplopia present, realignment surgery is regarded as a cosmetic procedure and therefore is not always covered by insurance (Marsh, 2014), but the American Association for Pediatric Ophthalmology and Strabismus (AAPOS) argues that strabismus surgery is reconstructive, not cosmetic, and that most insurance companies should cover the cost (American Association for Pediatric Ophthalmology and Strabismus, 2019). The appearance of strabismus is associated with a variety of negative psychosocial consequences. Self-reports from strabismus patients reflect difficulties in self-image, securing employment, relationships, work, school, and sports (Satterfield, Keltner, & Morrison, 1993). Children with strabismus score significantly lower on quality-of-life tests, particularly in categories such as physical, social, and educational functioning (Hatt, Leske, Kirgis, Bradley, & Holmes, 2007; Wen et al., 2011). This is also significant after controlling for gender, age, race, family income, other health conditions, and prior knowledge of the diagnosis, demonstrating that the social and physical impacts of strabismus are ubiquitous (Wen et al., 2011).

The presence of strabismus can elicit a strong social prejudice among children and adults alike. Children in digitally altered photographs that simulate esotropic and exotropic strabismus are rated negatively on intelligence, health, trustworthiness, being hardworking, happiness, cuteness, hesitancy, aggressiveness, activeness, and sentimentality compared to their unedited photos (Uretmen et al., 2003). Similarly, digitally altered photos of adults produced significantly negative perceptions of attentiveness, communication skills, competency, dependability, emotional stability, honesty, humor, intelligence, leadership ability, organizational skills, and sincerity compared to their unedited photos (Olitsky et al., 1999).

Intervention

Patients who have undergone successful corrective ocular-alignment surgery report significantly better self-esteem, self-confidence (Nelson, Gunton, Lasker, Nelson, & Drohan, 2008), physical function, general health, social function, mental health (Dickmann et al., 2013; Hatt, Leske, Liebermann, & Holmes, 2012), distance stereoacuity (Lal & Holmes, 2002; Yildirim et al., 1999), and binocular function (Mets, Beauchamp, & Haldi, 2004). However, the ocular alignment success rate of strabismus surgery is only about 70% (Hatt et al., 2012; Yildirim et al., 1999), and severe complications can occur during strabismus surgery, including perforation, slipped muscle, infection, scleritis, lost muscle, and retinal detachment. About 1 in 400 strabismus surgeries include one of these complications (Bradbury & Taylor, 2013). Successful surgeries come with a range of complications as well. Strabismus surgery does not reestablish balanced eye dominance, so while the ocular axes are realigned, ocular dominance imbalance remains (Zhou, Wang, Feng, Wang, & Hess, 2017). Additionally, body sway has been shown to increase significantly immediately following strabismus surgery (Legrand, Bui-Quoc, & Bucci, 2011; Matsuo, Narita, Senda, Hasebe, & Ohtsuki, 2006). The majority of behavioral-based interventions currently being used are in the form of simple eye exercises such as “pencil-pushups,” but their efficiency and retention rates are generally low and their use remains controversial (Rawstron, Burley, & Elder, 2005).

It is surprising that there is a lack of clinical trials for other behavioral-based oculomotor training paradigms (Jennings, 2000), as certain nonsurgical interventions for oculomotor deficiencies have been proven to be successful (Karlsson, 2017). Oculomotor auditory biofeedback training has been demonstrated to contribute to the ability to maintain ocular alignment (Goldrich, 1982) as well as fixational accuracy of eye movements (Hung, Ciuffreda, Carley, Fang, & Menditto, 1988), but these approaches have not been widely adopted. More recently, oculomotor training with visual feedback has been shown to significantly improve fixation stability and accuracy for peripheral vision tasks in healthy controls with simulated central vision scotomas, and this improvement is retained over the course of at least one week (Rose & Bex, 2017). If successful, oculomotor training paradigms for strabismus may allow the eyes to realign without invasive surgery or as postsurgical physiotherapy and may be used intermittently during follow-up, all of which may lead to lower rates of recidivism. Because of these factors, we propose a proof-of-concept investigation of a feedback-based oculomotor training paradigm for oculomotor control. We acknowledge that because we are using healthy controls in our paradigm, our findings cannot yet be directly generalized to the strabismic population. We hope, however, that this proof-of-concept study will lead to a clinical application that can be used for rehabilitation in people with strabismus.

Methods

This study was approved by the Institutional Review Board of Northeastern University and adhered to the tenets of the Declaration of Helsinki.

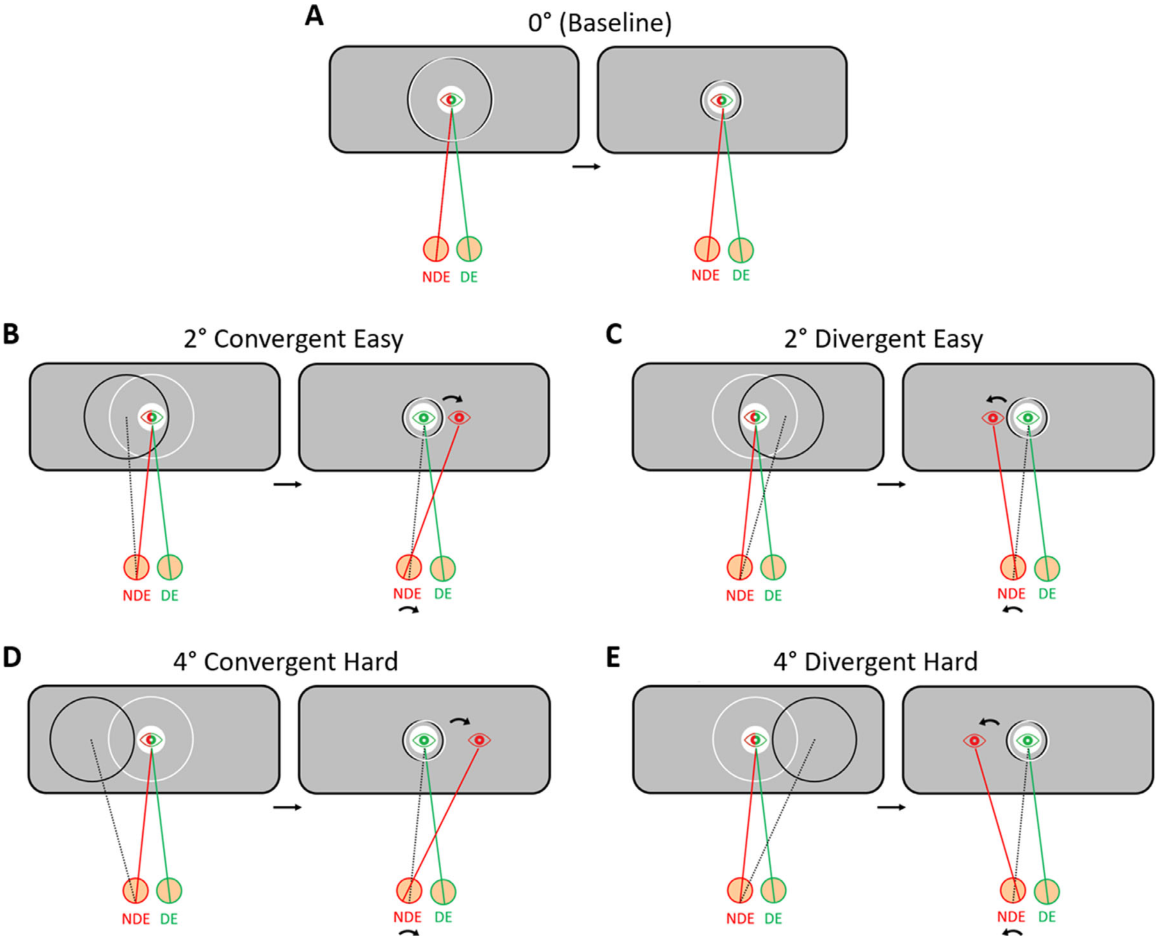

Figure 1. Schematic of the oculomotor training paradigm for a right eye dominant subject. Target and white ring were presented only to the dominant eye (DE, green), and black ring was presented only to the nondominant eye (NDE, red). The observer's task was to move their eyes to center the gaze-contingent rings on the target. The sizes of the rings decreased when the target was within the rings and decreased while the target was outside the rings. When introducing a vergence training offset, the center of the black ring was shifted away from the fovea in the nondominant eye so that when a participant was looking straight ahead, the white ring was centered with gaze and the black ring was shifted either temporally or nasally by the angle required for vergence training. Note that the angle that the black ring is shifted in each condition depicted above has been exaggerated for clarity. (A) Baseline, no offset. When no offset is applied, both rings are centered with gaze. When the rings overlap around the target, they are decreased in size to provide feedback for success. (B, D) Convergence training: the gaze-contingent ring in the NDE was offset temporally, 2° for easy (B) or 4° for hard (D), relative to the center of gaze, which required the observer to rotate their NDE nasally to center the black ring on the target dot. (C, E) Divergence training: the gaze-contingent ring in the NDE was offset nasally, 2° for easy (C) or 4° for hard (E), relative to the center of gaze, which required the observer to rotate their eye temporally to center the black ring on the target dot. Observers were required to achieve the trained vergence posture and hold it for at least 1 s before proceeding to the suppression test.

Design

We propose an oculomotor intervention technique designed to help patients with misaligned horizontal visual axis strabismus. We used normal-vision controls in this exploratory study to investigate the retention of learning and transfer of learning effects within this oculomotor alignment training paradigm. The dichoptic training paradigm is described in detail in the Procedure section below and in Caoli et al. (in press), and it is illustrated in Figure 1. In summary, the paradigm involves the dichoptic presentation of a static fixation target exclusively to the dominant eye and independent dynamic, gaze-contingent rings presented separately to the dominant and nondominant eyes. A gaze-contingent stimulus is one whose position on the screen is updated in real time such that the position of the retinal image relative to the fovea is constant and independent of eye movement. The observers' task is to move their eyes to control the positions of the gaze-contingent rings to center them on the fixation cross and hold this vergence posture for a minimum of 1 s. Example gaze position traces are provided in Caoli et al. (in press). Feedback is provided in real time by the size of the rings, which decrease in size during correct alignment or increase in size during incorrect alignment. Offsetting the center of the ring in the nondominant eye so it is no longer centered on the fovea requires the observer to make convergent or divergent eye deviations to perform the task, and the size of the offset specifies the magnitude of deviation. Shifting the center of the gaze-contingent ring temporally on the screen relative to that eye's gaze position (nasally on the retina relative to the fovea) requires convergent deviation to align the rings on the screen, as the participant must move his or her nondominant eye nasally to correct for the offset. Shifting the center of the gaze-contingent ring nasally requires divergent deviation, as the participant must move his or her nondominant eye temporally to correct for the offset. Note that throughout the task as well as upon successful completion of the task, a subject with normal retinal correspondence perceives two rings whose centers are offset by a constant amount.

Each condition was made up of two parts: vergence type (convergence or divergence) and vergence magnitude (2° or 4°). In convergence conditions, the center of the gaze-contingent ring was shifted temporally in the nondominant eye. This required the observer to deviate his or her nondominant eye nasally (convergent/esotropic) to align the ring on the target. In divergent conditions, the center of the gaze-contingent ring was shifted nasally in the nondominant eye. This required the observer to deviate his or her nondominant eye temporally (divergent/exotropic) to align the ring on the target. There were two levels of offset in each condition: ± 2°, labeled "easy," and ± 4°, labeled "hard." Subjects completed six training blocks, in two sessions of a baseline followed by three blocks, for different combinations of convergent/divergent and easy/hard conditions, as described below.

We were interested in a few key interactions. First, we wanted to see if there was a retention of learning in oculomotor training. In order to study this, we needed to run subjects twice, with each session being scheduled one week apart. Five subjects could not come exactly one week later due to scheduling conflicts. Two subjects came in 5 days after their first session, two subjects came in 8 days after their first session, and one subject came in 12 days after their first session. Subjects completed the same order of conditions in their second session as they did in their first session.

Second, we were interested to see if there was a transfer of learning effect both within conditions and between conditions.

Transfer of learning between vergence conditions is defined as the performance on one vergence condition given that a subject has been previously trained on a different vergence type. For example, would the *convergent easy* condition be improved after previous training on a *divergent* condition, compared to *convergent easy* as the first condition (Figure 2B)? This comparison was made in a between-subjects analysis by looking at the *divergent easy* and *convergent easy* conditions that were completed third and comparing them to performance on *divergent easy* and *convergent easy* conditions that were completed first. In both cases, a subject had never experienced the specific vergence condition before.

Transfer of learning within vergence conditions is defined as the performance on an offset level given that a subject has been previously trained on the same vergence type but at a different level. For example, would the *convergent hard* condition be improved after training on *convergent easy* condition, as compared to *convergent hard* as the first condition (Figure 2C)? This comparison was made in a between-subjects analysis by looking at all four conditions when they were completed second and comparing them to their respective conditions when they were completed first. In all cases, the subject had experienced the same vergence type before but at a different level.

For both transfer of learning types, we made comparisons using the Session 1 data only to avoid potential confounds of intervening conditions and time. We were interested in comparing performance for the first experience of a specific condition, and second session data would not be indicative of experiencing a condition for the first time.

For the purposes of this study, we determined that not every order combination of these four conditions was necessary to investigate retention and transfer of learning. Instead, we identified four total possible combinations that would be required in order to analyze the interactions that we were interested in (Figure 2A). We speculate that if successful, the most likely clinical application for this approach would be to correct large angle convergent or divergent deviations. We also speculate that since small angles are easier to correct than large angles (based on our pilot analyses and confirmed in the following data), transfer from small to large angle deviations may facilitate rehabilitation of resistant cases. We could not, however, identify a translational case to study transfer of learning between convergent and divergent hard conditions, and we therefore omitted this combination and excluded DH from the third-order position. There are no available estimates of variance on which to base a power calculation, and therefore we recruited 8 participants per condition for a total *N* of 32.

Binocular sensory assessment

There is a close relationship between ocular alignment and binocular vision (e.g., see Read, 2015), but surgical corrections of ocular deviation do not necessarily immediately correct binocular vision deficits

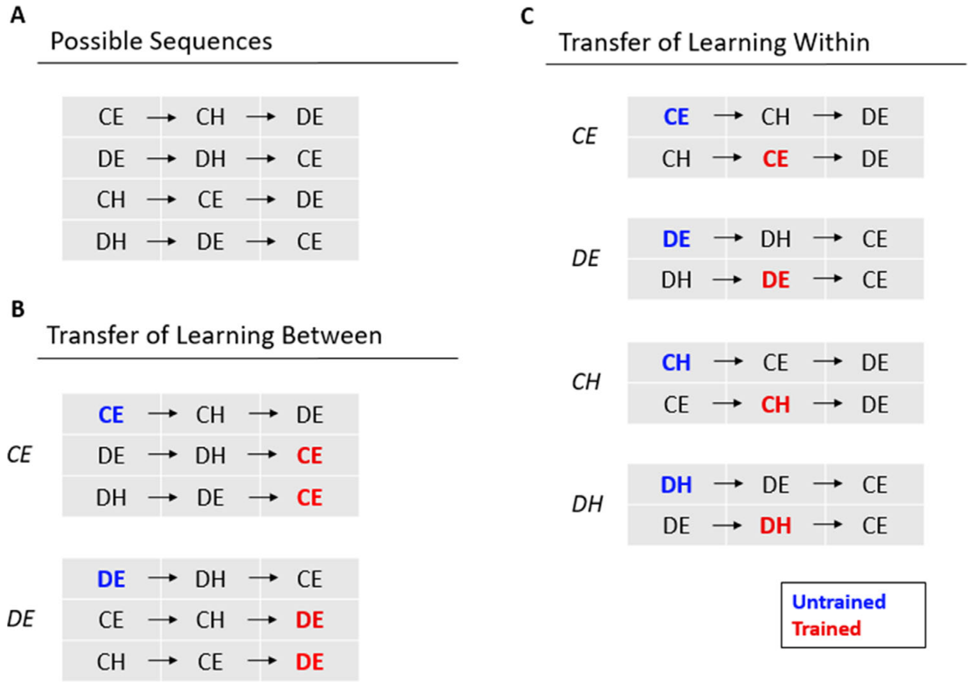

Figure 2. Subjects completed two sessions consisting of a baseline block followed by three training blocks of 40 trials each. In each block, subjects completed ocular alignment training for convergent easy (CE), convergent hard (CH), divergent easy (DE), or divergent hard (DH) deviations. (A) There were four possible condition sequences, distributed randomly across subjects. Each subject completed the same sequence in their first and second sessions. (B, C) List of the comparisons made for the transfer of learning between conditions and the transfer of learning within conditions. Data were compared only from the first training session. Comparisons in (B) identify the transfer of learning between conditions, and comparisons in (C) identify the transfer of learning within conditions. In all cases, untrained conditions are those that were completed first in a block, with no prior experience on any condition. Trained conditions for the between transfer are those completed third, where there had been training on a different vergence type beforehand (B). Trained conditions for the within block are those given second, where there has been training on the same vergence type beforehand (C), but at a different level.

(Zhou et al., 2017). We therefore sought to quantify whether the shifts in ocular alignment during the training were associated with binocular vision changes. Since ocular deviation disrupts interocular disparity, we elected not to attempt to measure stereoacuity. We therefore measured interocular suppression with a modified binocular balance paradigm (Kwon et al., 2015). The suppression stimulus was a Gabor patch with a peak spatial frequency of 4 c/deg and a standard deviation of 1.7°. The orientation of the Gabor was 45° in one eye and –45° in the other eye, at random, across trials. The observer's task was to indicate by pressing a corresponding keyboard arrow which orientation was dominant, guessing if necessary. The Michelson contrasts of the Gabors summed to 100%, with an initial value of 50% plus a uniform random deviate from ± 10%. Their contrasts were under the control of a QUEST staircase (Watson & Pelli, 1983), designed to converge on a contrast ratio that generated 50% responses for the Gabor in each eye. The proportions of left eye/right eye results for each contrast ratio were fit with a cumulative Gaussian distribution using MATLAB's *fit()* function, from which the balance point was estimated at the contrast, producing 50% responses for each eye.

Participants

In total, 39 subjects (16 male, 23 female) with self-reported normal or corrected vision participated in this study. A stereoacuity test was not performed, and participants wore corrective glasses or lenses as needed. The use of glasses or contacts was rarely disruptive during the study, but in two cases, eye-tracking calibration issues arose from the use of corrective lenses. These subjects were dismissed and their data were not used. Ocular dominance was determined with an ABC test (Miles, 1929) with a hole created by the subject's hands. Seven subjects were removed due to corrupted or mishandled data files ($n = 3$), eye-tracker calibration issues ($n = 2$), or failure to return for a second session ($n = 2$). These participants were eliminated as soon as issues were identified, and data collection continued until 32 usable subjects completed the study (13 male, 19 female). Participants were mainly recruited from the northeastern undergraduate population and received credit in an Introductory Psychology course in compensation for their time. Two subjects were recruited from the author's laboratories; these two subjects did not receive course credit and instead willingly offered their time for participation in the study. All subjects were naïve as of their first session, and all signed a consent form approved by the University Ethics Board. No subjects reported severe discomfort (such as diplopia or confusion) during or after the study, but two subjects informally reported mild eye strain or headache while attempting the harder tasks but not sufficient to withdraw from the study. Subjects were allowed to take breaks in between trials as needed and were not required to complete a block if uncomfortable (see Procedure).

Apparatus

Stimuli were presented on a 27-in. BenQ XL2720Z LCD monitor (BenQ Corporation, Taipei, Taiwan) set to a screen resolution of 1,920 × 1,080 pixels at 120 Hz and run using a Dell Optiplex 9020 desktop computer (Dell Inc. Round Rock, TX) with a Quadro K420 graphics card. The experiment was programmed and run using MATLAB (The MathWorks, Inc., Natick, MA) and the Psychophysics Toolbox Version 3 (Brainard, 1997; Pelli, 1997). Participants viewed the screen binocularly while wearing wired NVidia 3D Vision shutter glasses. Observers were seated at a distance of 53 cm from the monitor with head stabilization secured via chinrest. Eye movements were recorded using an SR Research Eyelink 1000 (SR Research Ltd. Mississauga, Ontario, Canada) and the MATLAB Eyelink Toolbox (Cornelissen et al., 2002). The sampling rate was set to 1,000 Hz. A 9-point calibration task was completed while wearing the shutter glasses before the start of each block.

Procedure

Participants were assigned to one of four condition sequences that were assigned to participant number randomly using a random number generator in MATLAB. All sequences began with a baseline task (0° offset) for 40 trials, followed by three condition sequences of CE, DE, CH, and DH as illustrated in Figure 2A. Before each condition, participants were given a manual demonstration of approximately where the convergence or divergence position would land in their real-world gaze. For example, CE would be similar to fixating a point a little in front of the screen, CH a point far in front of the screen, DE a point a little behind the screen, and DH a point far behind the screen. These instructions helped participants get a better conceptualization of approximately how much they needed to converge or diverge their eyes.

Each trial consisted of a white target dot (15 cd/m^2 through the shutter glasses) that was presented at one of nine locations at random within the central 15° of the display on a uniform gray background (7 cd/m^2 through the shutter glasses). The target dot was stationary and was not visible to the nondominant eye. A black, gaze-contingent ring was presented to the nondominant eye, while a white, gaze-contingent ring was presented to the dominant eye. The on-screen positions of the rings were independently updated according to the real-time gaze estimate of each eye. The diameter of the rings was initially 1.75°, and their size was continuously adjusted to provide feedback. The diameters decreased by 3.7 deg/s (1 pixel per frame) while the target was within the ring or increased at the same rate while the target was outside the ring. The observer's task was to align the gaze-contingent rings so that they were overlapping, centered on the target dot. During the baseline task (0° offset), the rings were already overlapping, as the center of each ring remained centered with gaze. For 2° and 4° of offset, the center of the black ring for the nondominant eye was shifted temporally (for convergent conditions) or nasally (for divergent conditions). Because the fixation target was presented only to the dominant eye, the dominant eye needed to remain mostly stable (but could move within a specified tolerance of 1°), while the nondominant eye would have to shift nasally (for convergent conditions) or temporally (for divergent conditions) in order to overlap the rings. If at any point the subject was unable to complete the alignment task, they were allowed to press the enter key, which terminated the trial and advanced to the next trial at a new test location.

On successful trials, once both rings were centered within 1° of the center of the target for at least 1 s, the target and rings were removed and were replaced by the suppression stimulus for 300 ms. The suppression Gabors were then removed and participants reported the orientation of the dominant Gabor. The observer's response initiated the next trial and the subject could withhold a response to rest.

Between blocks, participants could sit back from the chinrest while the next condition was prepared. Calibration was repeated at the start of each block. After the baseline task had been completed, participants began a deviation training condition corresponding to the predetermined order that was associated with their subject number.

Because converging and diverging the eyes is a difficult task, not all participants were able to complete all trials. A trial was aborted if a participant exceeded more than approximately 1 min on a single trial. This criterion was originally longer, but there became issues of subjects not finishing all conditions within the hour timeslot. Because of this, a maximum trial time

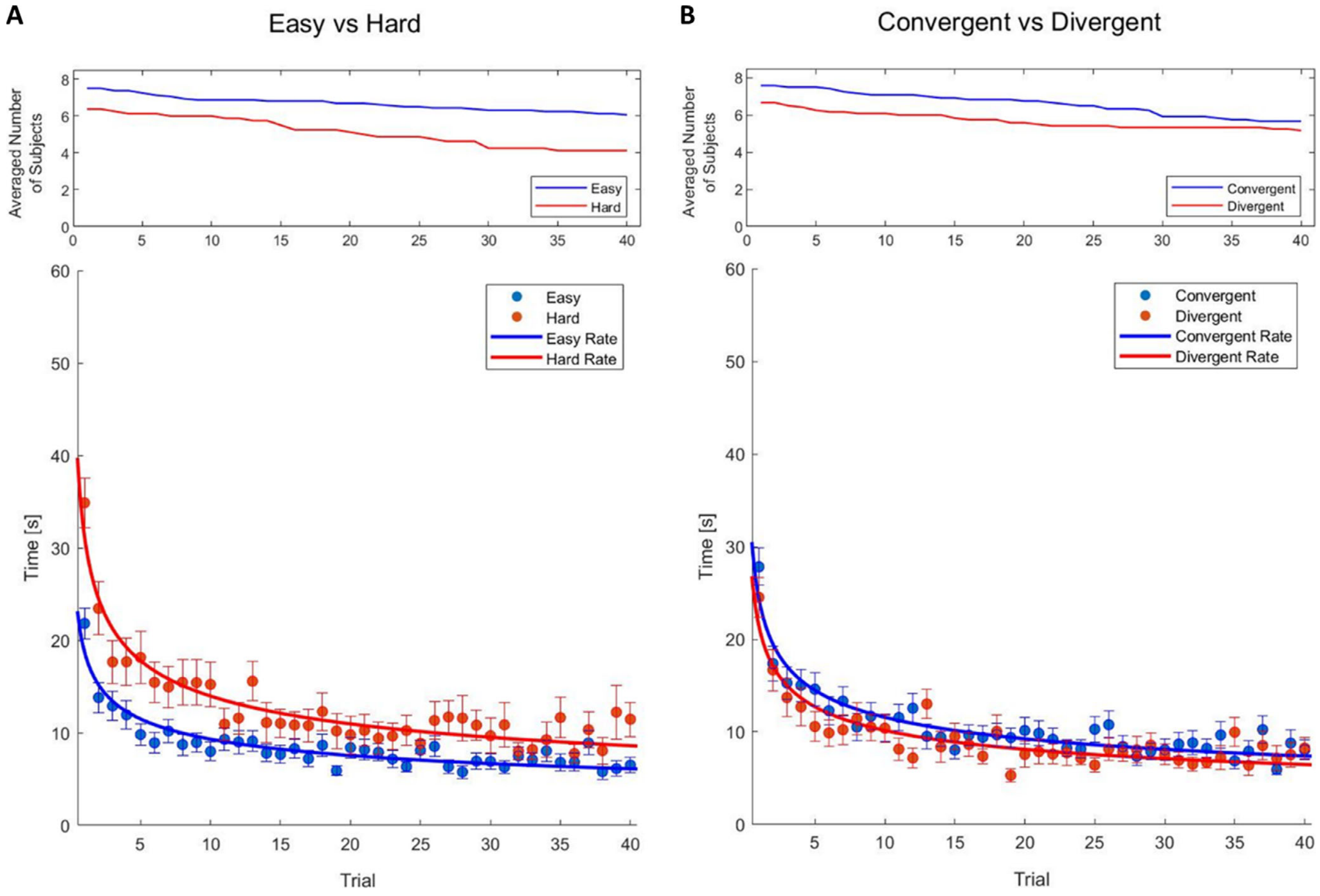

Figure 3. Mean trial times across subject (lower panels) and number of subjects completing trials (upper panels) for (A) all easy conditions (blue) compared to all hard conditions (red) and (B) for all convergent conditions (blue) compared to all divergent conditions (red). Error bars show ± 1 *SEM*; curves show the best-fitting power function.

of 1 min was set and during analyses, and trials that exceeded 1 min were capped at 1 min (in order to reduce the impact of outlier data from earlier participants). A condition was aborted if three trials were aborted in a row. A condition was also aborted when a total of 15 min was spent on a single condition. Times for aborted conditions were analyzed up until the trial at which the condition was aborted, and the number of trials completed was taken into account for each participant and condition.

Results

Overall difficulty of obtaining divergent and convergent gaze postures

Each subject completed up to 40 trials per condition and four conditions (baseline followed by three deviation conditions) in Session 1 and the same four conditions in the same order in Session 2, approximately one week later. First, we recorded the total number of trials successfully completed by each subject for each condition. Second, we recorded the time taken for each subject to achieve the training vergence posture on each trial.

Because we had a mixed design with unbalanced conditions, and because not every subject was able to complete all trials or conditions, we could not run an analysis of variance that equally matched all variables. In addition, a Shapiro-Wilk test confirmed that the data for each condition were not normally distributed and instead were always highly positively skewed (the majority of times centered on ∼10 s, with a higher spread above 10 s than below). We therefore determined that the most appropriate and conservative analysis was individual nonparametric *t* tests on our planned comparisons. Figure 3 shows the mean trial time across subjects and the mean number of subjects who completed each trial for all conditions. Analyses were conducted using the mean times across subjects for each trial. This was deemed more appropriate than using all times per subject because of covariance between trials. However, it should be noted that when running analyses using all times per condition instead

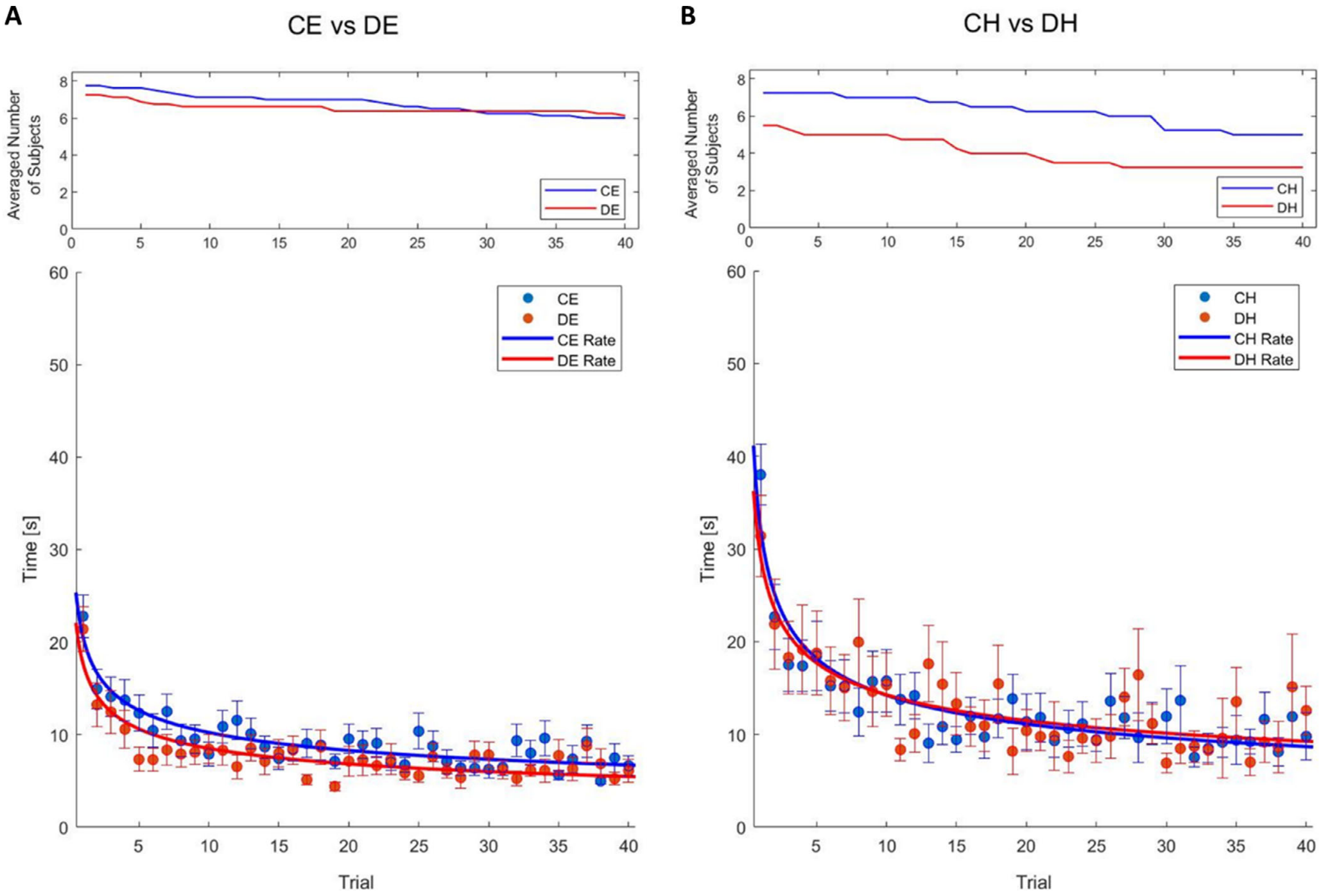

Figure 4. As Figure 3, except for convergent and divergent conditions at 2° (A) or 4° (B).

of averaged times per trial per condition, the main points we report here remain significant. Mean trial times were fit using the power series: $y = ax^b$ using MATLAB's *fit()* function with the model "power1." Error bars depict the standard error for each trial.

As expected, the time required to achieve the target vergence posture was significantly longer overall for larger angle hard conditions than for smaller angle easy conditions ($t = 10.533$, $p < 0.001$). Similarly, the total number of trials completed by subjects was significantly greater for easy conditions than for hard conditions ($t = -3.502$, $p < 0.001$).

Overall, the average time per trial was significantly shorter for divergence conditions than convergence conditions ($t = -5.029$, $p < 0.001$). However, subjects completed 15.94% more trials on convergence conditions than divergent conditions, but this difference was not significant ($t = -1.547$, $p = 0.190$).

Figure 4A shows the comparison between easy conditions for convergence and divergence, and Figure 4B shows the comparison for hard conditions, in the same format as Figure 3. For easy conditions, there was no significant difference between the number of trials completed for convergent and divergent ($t = -0.105$, $p = 0.917$) deviation conditions. However, the time taken to obtain the target vergence posture for divergent easy conditions was significantly lower than for convergent easy times ($t = -5.568$, $p < 0.001$). This suggests that convergent tasks at 2° are more difficult than divergent tasks at 2°. However, for hard conditions, subjects were able to complete significantly more convergent trials than divergent trials ($t = -2.375$, $p = 0.018$). In this case, the time taken to obtain the target vergence postures was not significantly different ($t = -1.107$, $p = 0.268$). This finding suggests that divergent tasks at 4° are more difficult than convergent tasks at 4°.

Retention

The lower panel in Figure 5 shows the overall time to complete trials in Session 1 and Session 2, averaged across all conditions. Overall, the time required to achieve the target vergence posture was significantly shorter in Session 2 than in Session one ($t = -2.695$, $p = 0.007$). Figure 5 (upper panel) shows the overall number of subjects who completed trials across all conditions, and significantly more trials were completed in Session 2 than in Session 1 ($t = 2.134$, $p = 0.033$).

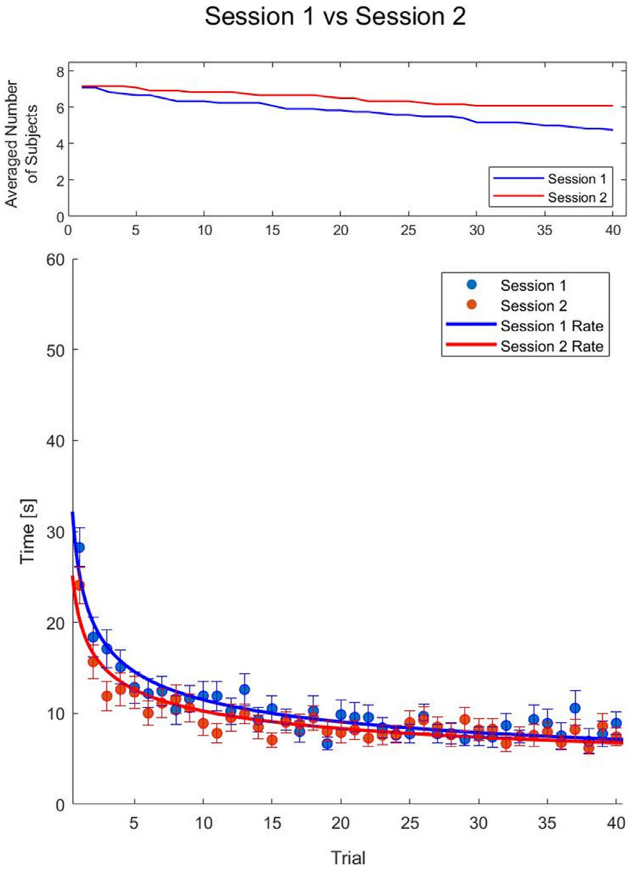

Figure 5. As Figures 3 and 4, the data show the average trial times and average number of subjects who completed each trial, combined across all conditions for Session 1 and Session 2.

These results suggest that some of the learning from Session 1 was retained approximately one week later.

Figure 6 illustrates how we quantified the proportion of learning retained, using values from the fitted power functions for each subject because these parameters incorporate all trial times across the duration of the block. In order to measure the amount of learning in Session 1 for each subject, we used the estimated durations of the first and last trials of Session 1 (F1 and L1, respectively). Next, to measure the proportion of learning retained for each subject, we used the estimated duration of the first trial in Session 2 (F2). This equation is as follows:

$$\frac{(\mathrm{F2} - \mathrm{F1})}{(\mathrm{L1} - \mathrm{F1})} \times 100.$$

In summary, to quantify the proportion of learning retained after training, we compared the difference in time from the end of the first session (after training completion) and the start of the second session (after one week of not training), normalized by the amount of learning in Session 1.

The data were well fit by a power function for 172 blocks, but the data from 36 of 172 did not follow a standard learning decay curve in either Session 1 or Session 2. In these cases, subjects required longer to obtain the target vergence posture at the end than at the start of the training session, and it is possible that these results reflect fatigue. These outlier curves strongly skewed the percentages of retention data, and we therefore use medians to analyze retention proportion data. The median percentage of learning retention is listed for each condition in Table 1. Overall, the median percentage of learning retained over one week is 53.9157%.

Transfer across task difficulty, within a deviation direction

Figure 7 shows the comparison of performance for CE, CH, DE, and DH conditions that were completed first (blue) or after subjects had completed CH, CE, DH, and DE conditions, respectively (red) (see Figure 2C). There was a significant increase in the number of trials completed for the DH condition for observers who had previously trained on the DE condition (+134.74%; $t = 2.165$, $p = 0.038$; Figure 7D, upper panel). However, there were no significant changes in the number of trials completed in the other conditions: The number of CH trials completed increased by 24.51% after CE training ($t = 0.506$, $p = 0.645$; Figure 7C, upper panel), the number of CE trials decreased 3.44% after CH training ($t = –0.181$, $p = 0.878$; Figure 7A, upper panel), and the number of DE trials decreased by 20.20% after DH training ($t = –0.767$, $p = 0.574$; Figure 7B, upper panel).

Based on the mean trial duration across subjects, the time taken to obtain the target vergence posture significantly improved for CE conditions after completing CH conditions ($t = –2.367$, $p = 0.018$; Figure 7A, lower panel), but there were no significant improvements for DE ($t = –5.87$, $p = 0.557$; Figure 7B, lower panel), CH ($t = 3.945$, $p < 0.001$; Figure 7C, lower panel), or DH ($t = .395$, $p = 0.693$; Figure 7D, lower panel). Note that the increases in time for CH and the nonsignificant decrease for DH conditions both come with the benefit of more trials completed, consistent with a survival trend.

We compared the transfer of learning models between and within conditions using an F statistic. As all models had the same number of parameters, the F statistic was computed as the ratio of the residual sum of squares for any two models being compared. Previous training significantly reduced the time required to obtain target vergence posture for DH conditions ($F(38, 38) = 1.18$ $p = 0.04$). However, the remaining three conditions—CE ($F(38, 38) = 1.2273$, $p = 0.23$), CH ($F(38, 38) = 1.39$, $p = 0.16$), and DE ($F(38, 38)$

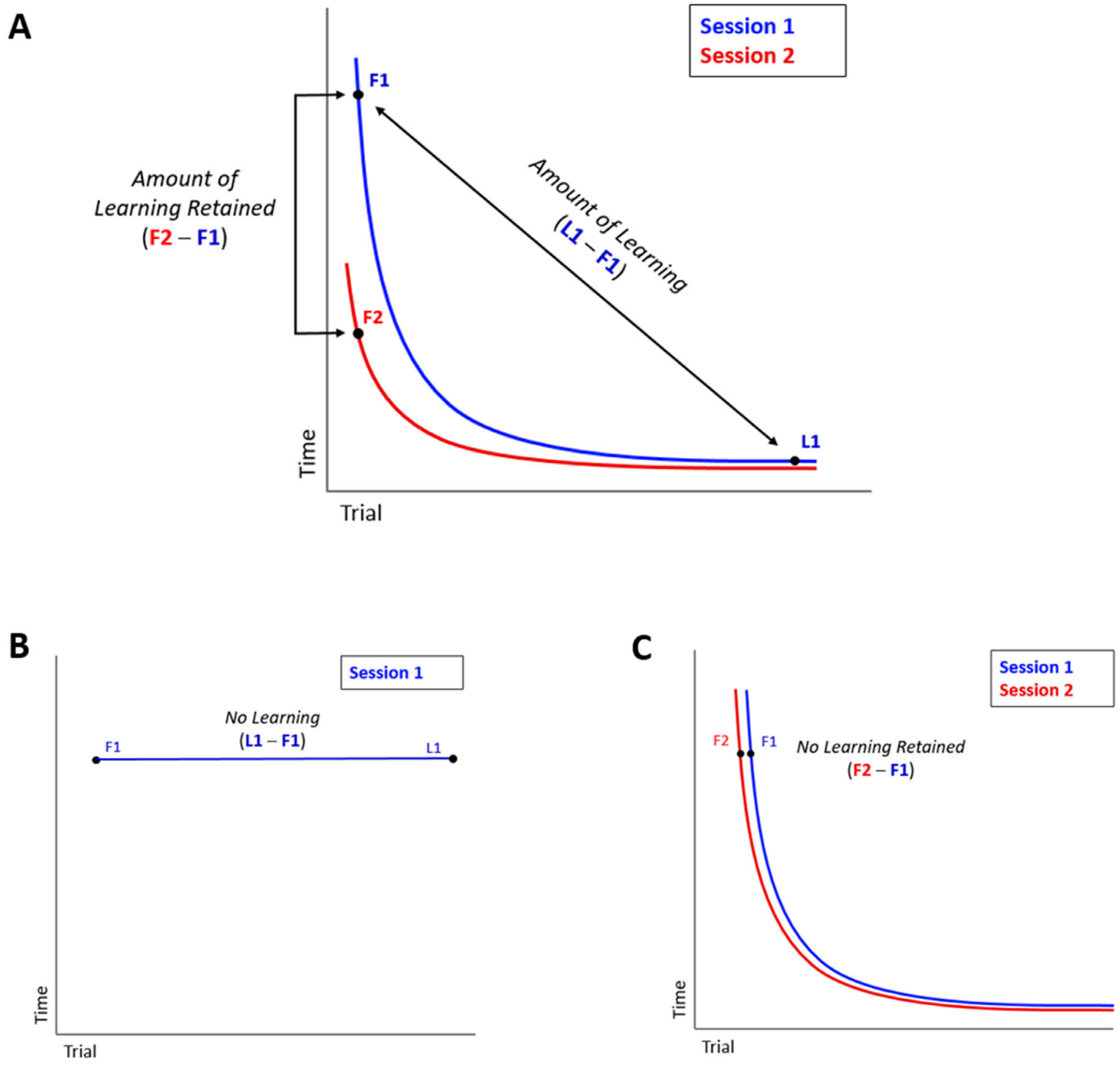

Figure 6. (A) Illustration of how the retention of learning was estimated. The amount of learning in Session 1 (blue curve) was estimated as the difference between the time taken to complete the first and last trials of Session 1 (L1 – F1). The amount of learning retained is the difference between the time taken to complete the first trial of Session 2 and the first trial of Session 1 (F2 – F1). The proportion of learning retained is the amount of learning retained divided by the amount of learning in Session 1. (B) Example of no learning during a session. (C) Example of no learning retained between sessions.

	Order			
	1	2	3 (E–H)	3 (H–E)
CE	86.89513122	83.43629765	13.29432248	89.9222934
CH	8.08954173	63.69913838		
DE	56.79123045	14.36103872	31.50297687	50.12851217
DH	51.04025257	63.36198447		

Table 1. Median proportions of learning retention for each condition.

$= 1.15$, $p = 0.33$)—did not show significant time improvements.

Transfer across deviation directions

Figure 8 shows the comparison of performance for CE and DE conditions that were completed first (blue) or after subjects had previously completed two divergent (DE and DH in either order) or two convergent (CE and CH in either order) conditions, respectively (red) (see Figure 2B). We combined the two sets of data for each sequence order (i.e., DE then DH with DH then DE; CE then CH with CH then CE) for the trained subjects.

Overall, the time taken to obtain the target vergence posture for DE was significantly faster for subjects who had previously trained on two convergent

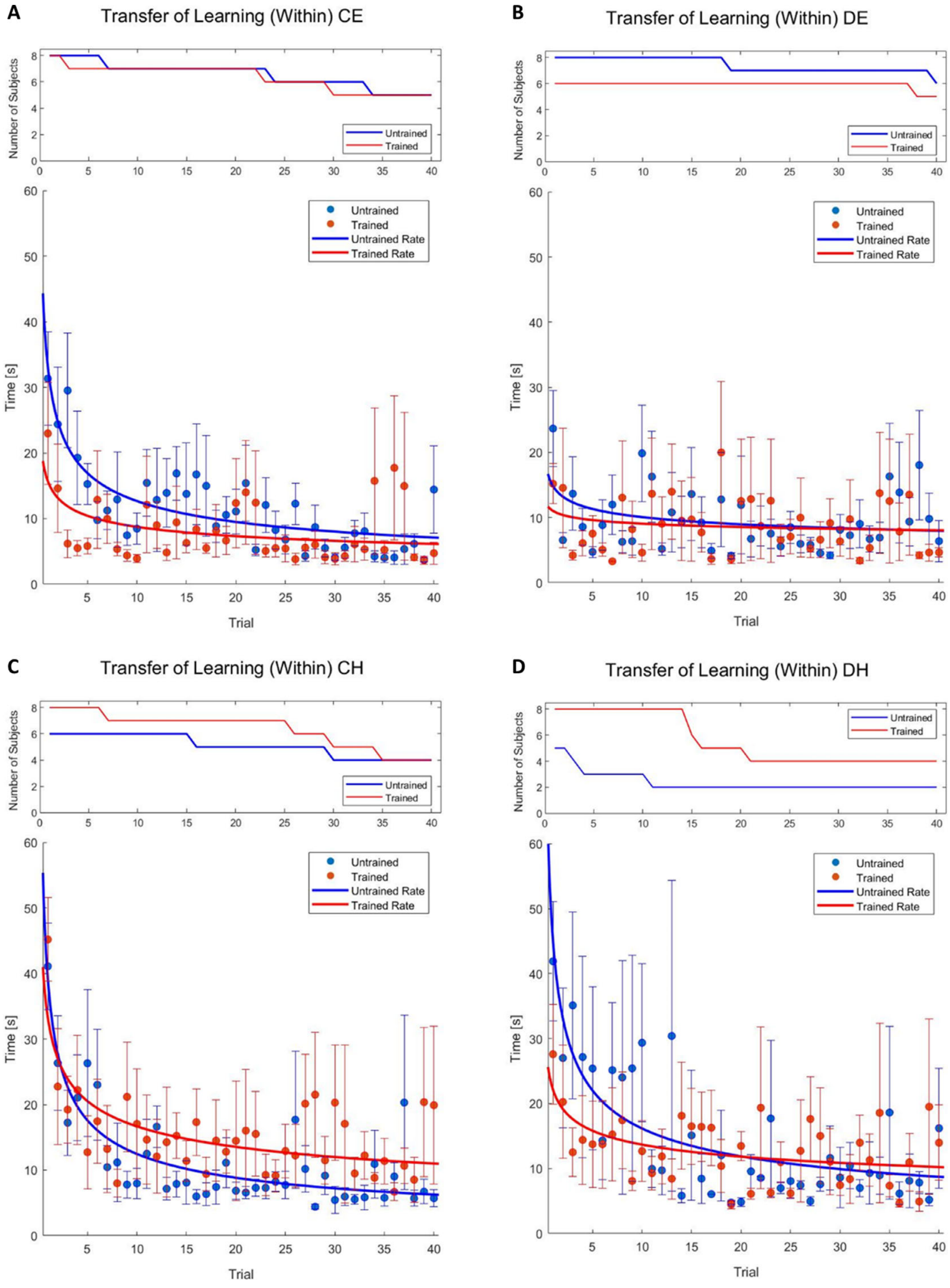

Figure 7. Transfer of learning within a deviation direction, across deviation magnitudes. The lower panels show the mean trial times and upper panels show the mean number of subjects who completed each trial for (A) convergent easy (CE), (B) divergent easy (DE), (C) convergent hard (CH), and (D) divergent hard (DH). Blue data show the results for the first condition in the sequence, and red data show results for the second condition in the sequence (see Figure 2C). Error bars show ± 1 *SEM*, and curves show the best-fitting power function.

conditions ($t = -4.320$, $p < 0.001$; Figure 8B, lower panel). However, there was no significant difference for convergent conditions after previous training on divergent conditions ($t = -0.896$, $p = 0.370$; Figure 8A, lower panel). Slightly more trials were completed for trained convergent conditions (+5.34%, $t = 0.610$, $p = 0.653$; Figure 8A, upper panel), and fewer trials were completed for trained divergent conditions

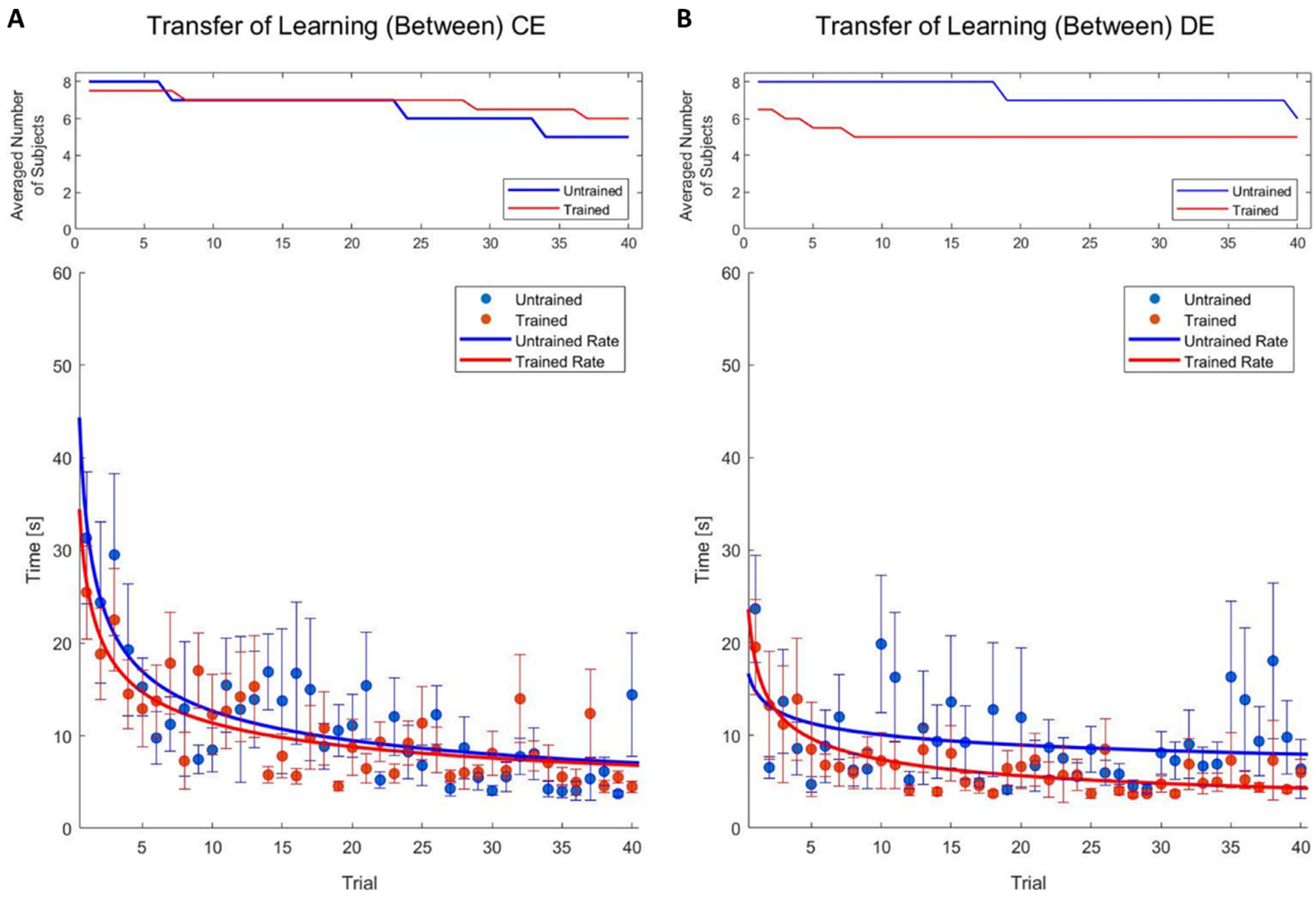

Figure 8. Transfer of learning across deviation directions. As Figure 7, except for easy conditions only after subjects had previously completed two conditions of training the other direction (see Figure 2B).

	Order			
	1	2	3 (E–H)	3 (H–E)
CE	$t = 2.768$ | $p = 0.028$*	$t = 0.666$ | $p = 0.527$	$t = –0.405$ | $p = 0.556$	$t = –1.68$ | $p = 0.136$
CH	$t = 0.454$ | $p = 0.669$	$t = –0.343$ | $p = 0.742$		
DE	$t = -1.374$ | $p = 0.212$	$t = –0.355$ | $p = 0.737$	$t = 1.548$ | $p = 0.166$	$t = –0.065$ | $p = 0.951$
DH	$t = 0.502$ | $p = 0.642$	$t = 0.368$ | $p = 0.724$		

Table 2. Results of significance testing for changes in interocular suppression of each individual condition compared to its baseline. Only data from Session 1 were used. Asterisk indicates the analysis has reached standard levels of significance.

(–30.30%, $t = –1.022$, $p = 0.417$; Figure 8B, upper panel).

When comparing the fit parameters, observers reached target vergence significantly faster when completing DE conditions after CE or CH training ($F(38, 38) = 6.89$, $p < 0.001$). A similar trend was found if the observer completed CE conditions after DE or DH training but did not reach statistical significance ($F(38, 38) = 1.60$, $p = 0.07$).

Suppression

Changes in interocular suppression can be found in Table 2. Our analysis of the contrast balance point for the dichoptic Gabor stimuli for each subject and each condition showed no overall significant change in interocular suppression with alignment changes. Changes in interocular suppression were measured by comparing the 50% report threshold of the nondominant eye in the baseline condition to the 50% report threshold of the nondominant eye in each individual condition for a single subject. Owing to unequal numbers of trials completed in each condition, we ran paired two-sample two-tailed t tests across all subjects within each given condition. The suppression data from some of the conditions were not normally distributed (DE-DH-CE, DH-DE-CE), and so nonparametric tests (Wilcoxon sign) were run for these conditions instead. The only uncorrected significant shift in suppression came from CE in the CE-CH-DE condition ($t = 2.768$, $p = 0.028$), which was not significant after Bonferroni correction for multiple comparisons ($p = 0.489$). All other comparisons resulted in a $p > 0.05$.

Discussion

We studied a real-time, dichoptic feedback-based method that modifies interocular alignment and aims to provide rehabilitation for strabismus. We previously showed that the method can be used to train reversible convergent and divergent gaze postures (Caoli et al., in press). In this proof-of-concept study, we studied the transfer and retention of oculomotor training in normally sighted control observers. The results showed transfer of training under some conditions: Training at one level generally transferred to another level within a gaze direction (convergence or divergence) but not between gaze directions. The results also showed retention of training over one week. The transfer results suggest that this dichoptic feedback approach may provide a scaffolding to train large angle strabismus by successive small steps, and the retention results suggest that these alignment benefits could be retained over time.

Subjects were required to use their eyes to align gaze-contingent rings on a target dot and to hold this gaze posture for at least 1 s. We recorded two classes of data; first, we measured the number of trials completed by each subject in each condition. Since this was a difficult task, not all observers completed every trial. Second, we recorded the time taken by each subject to obtain the target gaze posture each trial. In most cases, these two sources of data were in good agreement. However, the trial times and number of trials completed often indicated a different pattern; for example, there was a significant decrease in the time taken to complete CE after training but no significant change in the number of trials completed. These cases mainly occurred in the hard conditions, where subjects completed fewer trials on average. For example, only two out of eight subjects completed all 40 trials on their first divergent hard session. These two subjects had generally very fast times and did not struggle with the task. Most people, however, either could not complete the divergent task at all or had very difficult times completing only a handful of trials in their first attempt at this condition. This caused a rise in the number of aborted trials, leaving limited numbers of completed trial time data for a large number of trials for the majority of subjects. Consequently, the mean trial times become driven by the two participants who happened to be very good at the task. Compare this to the convergent hard condition, where more participants could complete a higher number of trials overall but at a higher average duration per trial. For these reasons, it is important to consider the average trial times with the overall number of trials completed per condition.

Overall, average trial times were shorter in Session 2 than in Session 1, and there were more trials completed for Session 2 than for Session 1, demonstrating that there is a retention of learning effect across conditions. Participants were better at the task (i.e., they completed more trials or completed them in a shorter time) when returning a second time, even though this second session took place about a week after the first session. Across all conditions, there was a positive median of a retention of learning effect. In other words, participants retained at least some of the learning that took place during their first session and as a result completed the first trial of their second session faster than the first trial of the first session while maintaining a consistent rate of learning.

All conditions had increases in the total number of trials completed from Session 1 to Session 2 (except for the third-order conditions, which were mixed). DH in the DH-DE-CE sequence is the only other condition in which there was a decrease in trials between Sessions 1 and 2, but this effect was mainly driven by one subject who was able to complete all 40 trials in their first session and 0 trials in their second session. Without this outlier, the increase in trials completed between Sessions 1 and 2 for the DH block of DH-DE-CE changes from a decrease of 15.8% to an increase of 45.5%. It is unclear why this subject was proficient in their first session yet was unable to complete any divergent hard trials in their second session. This is the only subject who had a drastic decline in trials completed between the two sessions.

Our results show that there may be an inequality of difficulty between the two tasks. Our null hypothesis was that 2° and 4° would be equally difficult in the two deviation directions, but this might not be the case. Our results suggest that diverging a small angle may be easier than converging a small angle, while diverging a large angle may be more difficult than converging a large angle. This is reflected in our observation that divergent easy times were significantly faster than convergent easy times, but there were significantly fewer trials completed for divergent hard than convergent hard conditions. In other words, the difficulty spike may be larger for divergent than convergent conditions, and future training protocols may need to factor in this difference to match the difficulty progression between the two vergence types. Our results showing transfer of learning across different levels within a deviation direction suggest that this can be accomplished by training participants incrementally on offsets from small angles to large angles for each vergence type, then correlating the average trial times and trials completed for each angle level between convergence and divergence.

There was no significant transfer of learning between divergent training conditions and subsequent convergent easy conditions. However, there was a transfer of learning effect for divergent easy conditions, where subjects who had previously trained on convergent tasks were significantly faster at completing divergent easy trials than those who were given divergent easy without any previous training. This effect is reflected as a significant decrease in time taken to achieve the target gaze posture, but these shorter trial times were associated with a 30.5% decrease in the number of trials completed compared to the untrained group. In other words, the untrained divergent group completed more trials than the trained divergent group, but they required more time to complete those trials. The average trial times for the trained group are therefore representative of less participants and as a result may not accurately reflect the performance of the group as a whole. This combination of results leads to the conclusion that there is no transfer of learning effect between conditions.

When looking at the transfer of learning within deviation directions, previous training on an easy condition helped for subsequent hard conditions, whereas previous training on hard conditions hurt for subsequent easy conditions. Training on a hard task before training on an easy one for the same deviation direction made the easy task harder to complete than if given the easy task first. This may be due to fatigue from attempting to accomplish the hard task at an earlier time in the same session that was necessary to make this comparison in the present study. However, it remains unknown whether such transfer of learning might occur on a more relaxed timescale, although the clinical translation for this combination (hard before easy) is less obvious than the other combination (easy then hard). On the other hand, when given an easy task before a hard task, the number of trials completed increased 24.51% for convergent conditions and 134.73% for divergent conditions. The average trial times for divergence were significantly shorter, and while the average trial times for convergence were significantly larger, the increase in trials completed reflects that of a survival trend.

There was no significant change in interocular suppression after subjects achieved a deviated ocular alignment, compared to the baseline with aligned eyes. Previous results have shown that strabismus surgery patients do not experience changes in binocular sensory vision immediately after their eyes are aligned (Zhou, Wang, Feng, Wang, & Hess, 2017). While there are significant differences between studies, Zhou et al. (2017) studied strabismus patients and we studied healthy controls; in both cases, changes in binocular motion function are not immediately accompanied by changes in binocular sensory function. There are two related implications: First, the suppression observed in strabismus is not observed with transient ocular misalignment in healthy subjects, and second, alignment interventions in strabismus patients may not be accompanied by immediate release from suppression. In the latter case, further perceptual therapy may be necessary. Taken together, these results suggest that rehabilitation for subjects with binocular sensory and motor deficits will require both sensory and motor interventions.

Overall, we found a retention of learning effect over a one-week period, and there was a transfer of learning demonstrated within the same deviation direction type going from easy to hard. This suggests that not only does training on a smaller angle of deviation improve performance on a larger angle of deviation but that this improvement is somewhat retained over the course of at least one week. We cannot say for sure if this paradigm will be as successful with strabismus patients, but we hope that this proof-of-concept study demonstrates the potential of our paradigm, and we are currently developing a clinical application that would allow us to test this definitively. Oculomotor training paradigms such as this may provide a noninvasive rehabilitation method for people with strabismus, with potentially lower recidivism rates and without the potentially negative outcomes that immediately follow surgery.

Keywords: strabismus, feedback, training, retention, learning

Acknowledgments

Supported by NIH R01 EY029713.

Commercial relationships: none.
Corresponding author: Kerri Walter.
Email: walter.ker@northeastern.edu.
Address: Northeastern University, Boston, MA, USA.

References

American Association for Pediatric Ophthalmology and Strabismus. (2019). Adult strabismus (eye misalignment). Retrieved from https://aapos.org/glossary/adult-strabismus

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436, doi:10.1163/156856897X00357.

Beauchamp, G. R., Black, B. C., Coats, D. K., Enzenauer, R. W., Hutchinson, A. K., & Saunders, R. A., …Felius, J. (2003). The management of strabismus in adults—I. Clinical characteristics and treatment. *Journal of AAPOS, 7*, 233–240, doi:10.1016/S1091-8531(03)00112–5.

Bradbury, J. A., & Taylor, R. H. (2013). Severe complications of strabismus surgery. *Journal of AAPOS, 17*, 59–63, doi:10.1016/j.jaapos.2012.10.016.

Caoli, A., Sabatini, S., Gibaldi, A., Maiello, G., Kosovicheva, A., & Bex, P. (in press). A dichoptic feedback-based oculomotor training method to manipulate interocular alignment. *Scientific Reports*.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers, 34*, 613–617, doi:10.3758/BF03195489.

Dickmann, A., Aliberti, S., Rebecchi, M. T., Aprile, I., Salerni, A., & Petroni, S., …Balestrazzi, E. (2013). Improved sensory status and quality-of-life measures in adult patients after strabismus surgery. *Journal of AAPOS, 17*, 25–28, doi:10.1016/j.jaapos.2012.09.017.

Epelbaum, M., Milleret, C., Buisseret, P., & Duffer, J. L. (1993). The sensitive period for strabismic amblyopia in humans. *Ophthalmology, 100*, 323–327, doi:10.1016/S0161-6420(13)32170–8.

Friedmann, L., Biedner, B., David, R., & Sachs, U. (1980). Screening for refractive errors, strabismus and other ocular anomalies from ages 6 months to 3 years. *Journal of Pediatric Ophthalmology and Strabismus, 17*, 315–317. PMID: 7441442.

Goldrich, S. G. (1982). Oculomotor biofeedback therapy for exotropia. *American Journal of Optometry & Physiological Optics, 59*, 306–317, doi:10.1097/00006324-198204000-00004.

Graham, P. A. (1974). Epidemiology of strabismus. *British Journal of Ophthalmology, 58*(3), 224–231, https://doi.org/10.1136/bjo.58.3.224.

Gunton, K. B., Wasserman, B. N., & Debenedictis, C. (2015). Strabismus. *Primary Care: Clinics in Office Practice, 42*, 393–407, doi:10.1016/j.pop.2015.05.006.

Hatt, S. R., Leske, D. A., Kirgis, P. A., Bradley, E. A., & Holmes, J. M. (2007). The effects of strabismus on quality of life in adults. *American Journal of Ophthalmology, 144*, 643–647, doi:10.1016/j.ajo.2007.06.032.

Hatt, S. R., Leske, D. A., Liebermann, L., & Holmes, J. M. (2012). Changes in health-related quality of life 1 year following strabismus surgery. *American Journal of Ophthalmology, 153*, 614–619, doi:10.1016/j.ajo.2011.10.001.

Harrold, A. L., & Grove, P. M. (2015). Binocular correspondence and the range of fusible horizontal disparities in the central visual field. *Journal of Vision, 15*(8), 12, doi:10.1167/15.8.12.

Hung, G. K., Ciuffreda, K. J., Carley, C. A., Fang, P., & Menditto, S. (1988). Auditory biofeedback to control vertical and horizontal eye movements in the dark. *Investigative Ophthalmology & Visual Science, 29*, 1860–1865.

Jennings, J. A. M. (2000) Behavioural optometry – a critical review. *Optometry Practice, 1*, 67–78.

Karlsson, V. (2017). Does nonsurgical treatment of exodeviations work? *American Orthoptic Journal, 59*, 18–25, doi:10.3368/aoj.59.1.18.

Kwon, M., Wiecek, E., Dakin, S. C., & Bex, P. J. (2015). Spatial-frequency dependent binocular imbalance in amblyopia. *Scientific Reports, 5*(1), https://doi.org/10.1038/srep17181.

Lal, G., & Holmes, J. M. (2002). Postoperative stereoacuity following realignment for chronic acquired strabismus in adults. *Journal of AAPOS, 6*, 233–237, doi:10.1067/mpa.2002.123399.

Legrand, A., Bui-Quoc, E., & Bucci, M. P. (2011). Re-alignment of the eyes, with prisms and with eye surgery, affects postural stability differently in children with strabismus. *Graefes Archive for Clinical and Experimental Ophthalmology, 250*, 849–855, doi:10.1007/s00417-011-1845-z.

Marsh, I. B. (2014). We need to pay heed to the psychosocial aspects of strabismus. *Eye, 29*, 238–240, doi:10.1038/eye.2014.283.

Matsuo, T., Narita, A., Senda, M., Hasebe, S., & Ohtsuki, H. (2006). Body sway increases immediately after strabismus surgery. *Acta Medica Okayama, 60*, 13–24, doi:10.18926/AMO/30754.

Mets, M. B., Beauchamp, C., & Haldi, B. A. (2004). Binocularity following surgical correction of strabismus in adults. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 8*, 435–438, doi:10.1016/j.jaapos.2004.07.003.

Miles, W. R. (1929). Ocular dominance demonstrated by unconscious sighting. *Journal of Experimental Psychology, 12*, 113–126, doi:10.1037/h0075694.

Nelson, B. A., Gunton, K. B., Lasker, J. N., Nelson, L. B., & Drohan, L. A. (2008). The psychosocial aspects of strabismus in teenagers and adults and the impact of surgical correction. *Journal of AAPOS, 12*(1), 72–76.e.1, doi:10.1016/j.jaapos.2007.08.006.

Olitsky, S. E., Sudesh, S., Graziano, A., Hamblen, J., Brooks, S. E., & Shaha, S. H. (1999). The negative psychosocial impact of strabismus in adults. *Journal of AAPOS, 3*, 209–211, doi:10.1016/S1091-8531(99)70004–2.

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

Pratt-Johnson, J. A., & Tillson, G. (1984). Suppression in strabismus—an update. *British Journal of Ophthalmology, 68*, 174–178, doi:10.1136/bjo.68.3.174.

Rawstron, J. A., Burley, C. D., & Elder, M. J. (2005). A systematic review of the applicability and efficacy of eye exercises. *Journal of Pediatric Ophthalmology and Strabismus, 42*, 82–88, doi:10.3928/01913913-20050301-02.

Read, J. C. A. (2015). Stereo vision and strabismus. *Eye (Basingstoke), 29*, 214–224, doi:10.1038/eye.2014.279.

Ritter, M. (1977). Effect of disparity and viewing distance on perceived depth. *Perception & Psychophysics, 22*, 400–407, doi:10.3758/bf03199707.

Rose, D., & Bex, P. (2017). Peripheral oculomotor training in individuals with healthy visual systems: Effects of training and training transfer. *Vision Research, 133*, 95–99, doi:10.1016/j.visres.2016.10.016.

Rutstein, R. P., Martin Cogen, P. S., Cotter, S. A., Kent Daum, O. M., Mozlin, R. L., & Julie Ryan, O. M., …Stephen Miller, O. C. (2010). Optometric clinical practice guideline care of the patient with strabismus: Esotropia and exotropia. In R. P. Rutstein (Ed.), *Reference guide for clinicians prepared by the American Optometric Association Consensus Panel on Care of the Patient with Strabismus*. American Optometric Association. 243 N. Lindbergh Blvd., St. Louis, MO.

Satterfield, D., Keltner, J. L., & Morrison, T. L. (1993). Psychosocial aspects of strabismus study. *Archives of Ophthalmology, 111*(8), 1100, doi:10.1001/archopht.1993.01090080096024.

Stidwill, D. (1997). Epidemiology of strabismus. *Ophthalmic and Physiological Optics, 17*(6), 536–539, https://doi.org/10.1111/j.1475-1313.1997.tb00094.x.

Uretmen, O., Egrilmez, S., Kose, S., Pamukcu, K., Akkin, C., & Palamar, M. (2003). Negative social bias against children with strabismus. *Acta Ophthalmologica Scandinavica, 81*, 138–142, doi:10.1034/j.1600-0420.2003.00024.x.

Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113–120, https://doi.org/10.3758/bf03202828.

Wen, G., Mckean-Cowdin, R., Varma, R., Tarczy-Hornoch, K., Cotter, S. A., Borchert, M., … Azen, S. (2011). General health-related quality of life in preschool children with strabismus or amblyopia. *Ophthalmology, 118*, 574–580, doi:10.1016/j.ophtha.2010.06.039.

Yildirim, C., Mehmet Mutlu, F., Chen, Y., & Altinsoy, H. I. (1999). Assessment of central and peripheral fusion and near and distance stereoacuity in intermittent exotropic patients before and after strabismus surgery. *American Journal of Ophthalmology, 128*(2), 222–230, doi:10.1016/S0002-9394(99)00079-3.

Zhou, J., Wang, Y., Feng, L., Wang, J., & Hess, R. F. (2017). Straightening the eyes doesn't rebalance the brain. *Frontiers in Human Neuroscience, 11*, 453, doi:10.3389/fnhum.2017.00453.