A comparison of therapies using Eyetronix Flicker Glass and standard adhesive patches in children with anisometropic amblyopia: A randomized controlled trial

A Preprint

Seung Hyun Min *
McGill Vision Research
McGill University

Shijia Chen †
Department of Ophthalmology
Wenzhou Medical University

Jinling Xu
Department of Ophthalmology
Wenzhou Medical University

Bingzhen Chen
Department of Ophthalmology
Wenzhou Medical University

Hui Chen
Department of Ophthalmology
Wenzhou Medical University

Yuwen Wang
Department of Ophthalmology
Wenzhou Medical University

Jiawei Zhou
Department of Ophthalmology
Wenzhou Medical University
zhoujw@mail.eye.ac.cn

Xudong Yu
Department of Ophthalmology
Wenzhou Medical University
yxd@mail.eye.ac.cn

July 21, 2020

Abstract

Objective: Eyetronix Flicker Glass (EFG) has been introduced an alternative treatment for amblyopia. It is based on the premise that amblyopia can better be treated by manipulating the visual input that enter both the normal and amblyopic eye. In contrast, only the normal eye gets deprived during the traditional patching therapy. We conducted a randomized clinical trial to compare the efficacy of the therapies using the Eyetronix Flicker Glass and standard adhesive patches in children with anisometropic amblyopia.

Design: A prospective, randomized controlled trial Participants: 31 children aged 4-13 years with anisometropic amblyopia Intervention: The patients were assigned two treatment groups and were treated for 12 weeks. Those in the first group were treated with the Eyetronix Flicker Glass for one hour per day, whereas those in the latter group were treated with the standard patches for two hours per day.

Main outcome measures: Best-corrected visual acuity and contrast sensitivity of the amblyopic eye, stereopsis and fusion range were measured at both before the treatment and 3, 6 and 12 weeks after the first day of the treatment.

*Also affiliated with the Department of Ophthalmology at Wenzhou Medical University
†co-first author with SHM

Results: Visual acuity of the amblyopic eye significantly improved in both groups after the treatment ($p < 0.05$), albeit more so in patients who had undergone the standard patching therapy. A significant improvement in contrast sensitivity of the amblyopic eye at 3, 6 and 12 cpd was also observed in patients who had undergone the standard patching therapy ($p < 0.05$) but not the EFG therapy ($p > 0.05$). However, no improvement in stereopsis and fusion range was found in patients from both treatment groups.

Conclusion: Both therapies improved monocular visual functions, such as visual acuity and contrast sensitivity. However, the improvement was significantly larger after the standard patching therapy than that of the EFG therapy. However, both therapies did not improve fusion range and stereopsis. Our results indicate that although the EFG therapy deprives visual input in a binocular fashion, it does not improve binocular functions in the amblyopic population.

***K*eywords** Amblyopia · Randomized clinical trial · Eyetronix Flicker Glass · Monocular occlusion · Children

1 Introduction

Amblyopia is a neurodevelopmental disorder that has numerous visual consequences such as poor visual acuity, contrast sensitivity and stereopsis [1]. It arises from abnormal visual experience during the critical period when the brain function is most susceptible to changes in visual input [2, 3]. Hence, rewiring of the visual cortex and, hence, attaining normal visual experience for the visually impaired must take place before the critical period terminates. For the past 250 years, physicians have deprived the normal eye of amblyopic children for two to six hours for over the course of months to years [4]. The traditional patching therapy improves the visual acuity of the amblyopic eye by more than 2 lines of logMAR in over 50% of cases [5]. However, it has been less than ideal for several reasons. First, it discomforts the children as it enforces them to only see through the amblyopic eye [6, 7]. Second, it has a low compliance rate due to the discomfort [8]. Moreover, it has been associated with adverse psychosocial factors such as emotional distress [9]. Furthermore, the improved visual acuity of the amblyopic eye does not ensure an improvement in binocular vision [10, 11], which is more important for everyday visual experience and self-perception [12]. Traditional patching therapy has also been prescribed with pharmacological intervention, such as the injection of atropine into the fellow eye. However, the possibility for the complication from its toxicity is not negligible [13]. Therefore, for these reasons that illustrate the fact that the traditional patching therapy is not flawless, an alternative treatment for amblyopia that can bring about both long-lasting monocular and binocular visual improvements is needed.

Recently, liquid crystal glasses (LCG) have been introduced as a device to include in an alternative therapy to standard patching [14]. They have the structure of typical glasses, but their lenses are built with liquid crystal cells that enable the lenses to be either transparent or opaque [15]. The alternation between the transparent and opaque states can be extremely rapid and appears as flicker. As is the case in the standard patching therapy, a therapy that incorporates the LCG only deprives the normal eye and, hence, can be categorized as a monocular therapy. Therefore, the lens in front of the normal eye flickers. However, unlike traditional patching which entirely deprives the normal eye throughout the treatment period, LCG therapy still allows the children to see the world through both of their eyes because the lens in front of the normal eye alternates between being opaque and transparent. LCG therapy seems to be promising. Studies have shown that it improves the visual acuity of the amblyopic eye (i.e., monocular visual function), compliance and satisfaction of patients [14, 16, 17]. In sum, LCG therapy exemplifies the recent effort of researchers and clinicians in designing novel and efficacious treatment for amblyopia.

However, amblyopia has been recently understood as a complex and multifaceted disorder as it is characterized by not just poor monocular visual functions of the amblyopic eye but also imbalanced binocular interactions between the normal and amblyopic eyes [18, 19, 20, 21, 22, 23]. In normal observers, the two eyes suppress each other during normal viewing in a balanced fashion. However, in amblyopes, the suppression of the amblyopic eye on the fellow eye is slightly weaker than that of the fellow eye on the amblyopic eye [24, 23]. This results in a net imbalance resulting in

total dominance by the fellow eye. Therefore, an important step to treat amblyopia based on this abnormal binocular interaction is to redress this imbalance between the eyes. The disrupted binocular balance in amblyopia might be the reason why the traditional patching therapy, which only manipulate the visual input through one eye (i.e., normal eye), can be inadequate for a full recovery [10, 11]. Conversely, a binocular treatment, which manipulates input in both eyes, could be more beneficial. Researchers have already explored this possibility. For example, Schor, Terrell and Peterson (1976) demonstrated that flicker that alternately occurs in lenses for both eyes (i.e., alternate flicker) at about 2 and 7 Hz reduces an enhanced masking by the fellow eye on the amblyopic eye [25]. In other words, the formerly imbalanced binocular interaction became more balanced in amblyopic patients, and the visual acuity of the amblyopic eye improved. This study shows that binocular treatment strategy can benefit amblyopic patients.

Likewise, Eyetronix® have introduced Eyetronix Flicker Glass (EFG) on the principle that amblyopia can be better treated via binocular treatment. In a therapy that incorporates EFG, the lenses in front of the amblyopic and normal eyes are alternatively transparent and opaque; this is similar to alternate flicker as seen in Schor et al. (1976) [25]. Hence, EFG therapy deprives inputs from both eyes in an alternate fashion unlike the LCG therapy which deprives the normal eye only [17, 26]. If it is true that binocular therapy can improve binocular functions in amblyopia, as it was reported by Schor et al. (1976), the EFG therapy might benefit binocular functions such as stereopsis and fusion. Indeed, a recent study shows that EFG therapy for 12 weeks improves both the visual acuity of the amblyopic eye and stereopsis in 23 patients with anisometropia amblyopia [27]. Furthermore, a subset of the patients showed a sustained improvement in visual acuity and stereopsis for 12 weeks after the treatment. However, although Vera-Diaz et al. observed a noticeable improvement in visual acuity and stereopsis using the EFG therapy, they did not compare the efficacy with that of the standard patching therapy [27].

In this study, we performed a randomized controlled trial for the first time to compare the efficacy of traditional patching and EFG therapies in amblyopic children. We measured a wide range of both monocular and binocular visual functions (see Methods for details) before and during the treatment for over the course of 12 weeks. We show that the EFG therapy improves the visual acuity of the amblyopic eye and, to a lesser degree, its contrast sensitivity. However, standard patching therapy improved visual acuity and contrast sensitivity of the amblyopic eye more so than EFG therapy. Both therapies did not improve fusion range and stereopsis. In conclusion, our results indicate that although the Eyetronix Flicker Glass is binocular treatment, it does not improve binocular functions in the amblyopic population. Moreover, the standard patching therapy seems to be more effective in recovering monocular visual functions.

2 Materials and Methods

This randomized clinical trial is listed in the Chinese Clinical Trial Registry (http://www.chictr.org.cn) with the identifier ChiCTR2000034436. The Full trial can be accessed from the Chinese Clinical Trial Registry's website: http://www.chictr.org.cn/showproj.aspx?proj=56030. The research protocol and the informed consent forms were reviewed and approved by the ethics committee of the Affiliated Eye hospital at Wenzhou Medical University (2016-18-Q-11). All data were collected at the Affiliated Eye Hospital of Wenzhou Medical University.

2.1 Eyetronix Flicker Glass

Eyetronix Flicker Glass (http://eyetronix.com) utilizes liquid crystal technology. Liquid crystal has properties of liquid and a solid crystal. Its liquid properties enable the molecules within to readily rotate and move and make liquid crystal an excellent conductor of light.

Each lens of the EFG contains liquid crystal shutter cells [15], which are suspended in gel-like liquid (light blue space in Figure 1A). The gel-like liquid is sandwiched between one thin glass plate on one side and another plate on the other side (thin green structures in Figure 1A). On the outer side of each thin glass plates lies a polarizer film that controls the transmission of light (striped or dotted structure in Figure 1A). Light transmission through the lenses depends on the

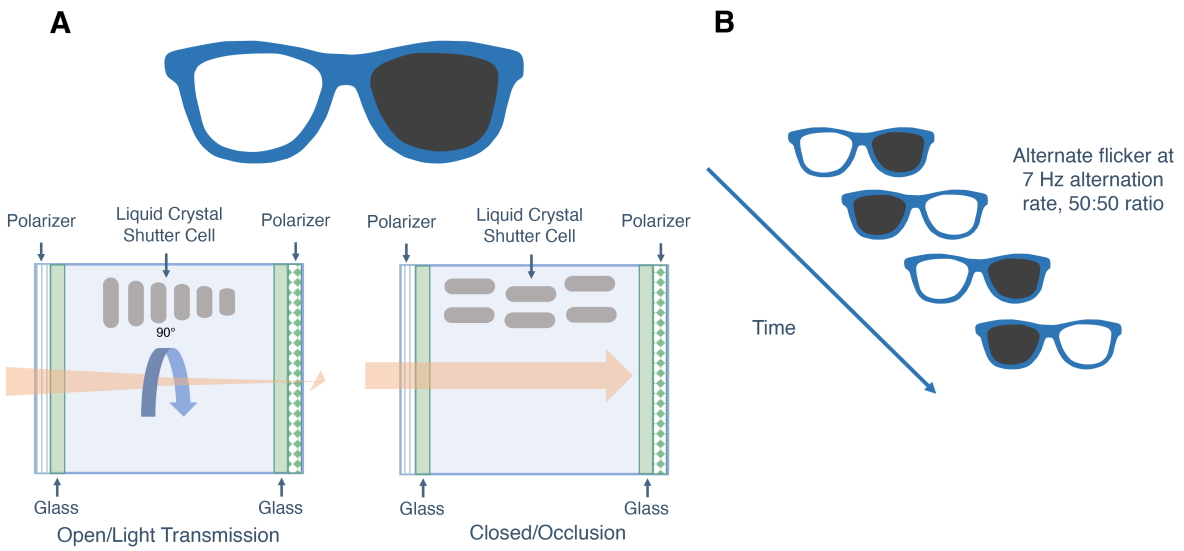

Figure 1: **A diagrammatic representation of the Eyetronix Flicker Glass.** A) The lens of the EFG can assume either a transparent or opaque state depending on whether the lens allows light transmission or not. When electric field is present, the lens assumes an 'open' state and enable light transmission. This will make the lens transparent. When electric field is absent, the lens assumes an 'closed' state and obstruct light transmission. This will make the lens opaque. B) An illustration of alternate flicker of the EFG. We programmed the EFG so that it would perform alternate flicker at 7 Hz alternation rate, 50:50 ratio between the lens in front of the amblyopic eye and that in front of the normal eye. Therefore, 1-hour treatment with the EFG would be equivalent to 30 minutes deprivation of the normal eye.

polarizer film. Since there are two glass plates in each lens, there are two polarizer films per lens. These two polarizer films are at perpendicular orientation to one another (90). The first polarizer only allows the horizontally oriented light to pass, and the second polarizer only the vertically oriented light because the orientations of the polarizers are perpendicular to one another (the striped pattern represents a perpendicular orientation to the dotted pattern in Figure 1A). Therefore, the light transmitted by the first polarizer will not be transmitted by the second polarizer in the lens of the EFG is in its natural state. When no external energy is applied, the lens is therefore opaque, and no light gets through (Closed/Occlusion state in Figure 1A).

However, by applying external energy, one can make the lens to be transparent (Open/Light transmission state in Figure 1A). When an electric field is present, the formerly perpendicular orientations of the polarizer films become parallel and aligned because the liquid crystal shutter cells (shown in grey structures in Figure 1A), which are suspended in gel-like liquid, will change their position. As a result, the light will be able to be transmitted through the lens. In other words, by applying electric field, one can control make the formerly opaque liquid crystal lens to become transparent within milliseconds (see Figure 1B).

EFG and Liquid Crystal Glasses (LCG) both utilize the liquid crystal technology and share the goal of replacing the standard patching therapy. However, they are slightly different. The LCG therapy only deprives the normal eye by only allowing the lens in front of said eye to shift between its transparent and opaque state. On the other hand, the EFG therapy occludes both the normal and amblyopic eyes in an alternate fashion (see Figure 1B) at a specific ratio and temporal frequency.

2.2 Patients

32 children (aged 4-13) with untreated, moderate to severe unilateral and anisometropic amblyopia participated in our study. One of them received another treatment during the study and had to be excluded, so 31 children completed the study. The clinical details of the patients are provided in Table 1. The sample size was determined based on the previous studies of EFG and LCG [17, 27]. We obtained informed consent from their parent or legal guardians. Their guardians volunteered to participate in the study and accepted the random assignment of each patient to either Group 1 (traditional patching therapy) or Group 2 (therapy with the EFG).

Eligibility inclusion criteria: 1) Age range of 4 to 13 years. 2) Definite genetic diagnosis of anisometropic amblyopia, best-corrected visual acuity (BCVA) of the amblyopic eye equal to or worse than 0.3 (logMAR) in 3–5 years age range and 0.15 (logMAR) in children over 6 years old, at least a two-line BVCA difference between the two eyes (according to the diagnostic criteria for amblyopia established by the Ophthalmology branch of the Chinese Medical Association in 2011). 3) Myopia of no more than -6.00 diopters (D), hypermetropia of no more than +9.00 D, astigmatism of less than 3.00 D, anisometropia of at least 1.50 D spherical equivalent or at least 1.00 D cylindrical equivalent. 4) Strabismus of no more than 20 prism diopters (Δ) according to prism and alternate cover test (PACT). 5) Better than a 0.7 logMAR BVCA of the amblyopic eye.

Eligibility exclusion criteria: 1) Ocular diseases, such as ptosis, refractive media opacity, fundus disease and optic neuropathy. 2) History of interocular or refractive surgery that affects vision. 3) History of treatment for amblyopia in the last three months before screening (except spectacle frames). 4) Photosensitivity epilepsy. 5) Confirmed or suspected conjunctivitis. 6) Allergy or intolerance to the test equipment or patch. 7) History of pharmacological intervention that may affect vision such as atropine. 8) Participation in another clinical study/trial within one month before the enrollment of the study.

2.3 Treatment procedure

After confirming for the eligibility of each patient, we randomly assigned all participants to one of the two treatment groups using a random number generator so that the group assignment was balanced for both groups (see Figure 2). EFG therapy intervention group (Flicker Group): In this treatment group, patients were instructed to wear the EFG rather than a standard adhesive patch. Schor et al. employed a rate for alternate flicker at 7 Hz [25]. Likewise, we used parameters of 7Hz and 50% duty cycle in this study. At 7 Hz, the lens assumes each state of transparent (open) and opaque (closed) state for 0.71+ seconds. At a duty cycle of 50% (50:50 ratio), EFG deprives the normal eye 50% of the time throughout the treatment. Therefore, 1-hour treatment with the EFG would be equivalent to 30 minutes deprivation of said eye (50%). The patients were asked to wear the EFG for one hour per day, 7 days per week throughout the 12-week treatment period. The EFG automatically recorded the duration and time in which the patients wore them.

Traditional patching control group (Patching Group): Patients were instructed to wear a standard, latex-free and adhesive style patch in front of their normal eyes for two hours, 7 days per week throughout the 12-week treatment period.

For all of the patients in Flicker Group and Patching Group, the flicker and patching treatment was combined with optimal refractive correction. The parents of the patients were given a form by which they reported a completion of the patching session each day throughout the treatment.

During the period of wearing the EFG or the patch, patients performed daily activities such as doing homework or watching television. During their first visit (Day 1), guardians of the patients within the Flicker Group were informed about the specifics of the EFG device, such as recharging and handling. Follow-up visits were scheduled after 3, 6 and 12 weeks from their first visit. During each visit, monocular and binocular visual functions were assessed. We had decided to follow up patients for 12 weeks because the investigators that first reported the benefit of the EFG therapy also followed up their patients for 12 weeks.

Table 1: **Baseline characteristics of the patients**

	Flicker Group		Patching Group	
	N	%	N	%
Gender				
Female	6	40%	8	50%
Age at enrollment (years)				
4 to <6	9	60%	6	38%
6 to <8	5	33%	5	31%
8 to 13	1	7%	5	31%
Mean ± SE	5.27 ± 0.28		6.38 ± 0.61	
Distance visual acuity of amblyopic eye (logMAR)				
0.15 to <0.2	1	7%	1	6%
0.2 to <0.3	3	20%	3	19%
0.3 to <0.4	3	20%	5	31%
0.4 to <0.5	3	20%	2	13%
0.5 to <0.6	3	20%	2	13%
0.6 to <0.7	1	7%	0	0%
0.7	1	7%	3	19%
Mean ± SE	0.39 ± 0.04		0.39 ± 0.05	
Distance visual acuity of fellow eye (logMAR)				
Mean ± SE	0.03 ± 0.01		0.01 ± 0.02	
Interocular acuity difference (logMAR)				
Mean ± SE	0.37 ± 0.05		0.38 ± 0.04	
Baseline stereoacuity (arc seconds)				
Nil (converted to 4500)	4	27%	6	40%
800	0	0%	1	7%
400	3	20%	2	13%
200	2	13%	0	0%
140	0	0%	2	13%
100	1	7%	0	0%
80	0	0%	3	20%
60	1	7%	1	7%
50	2	13%	0	0%
40	2	13%	1	7%
Mean ± SE (arc seconds)	1329.33 ± 512.11		1826.25 ± 536.81	
Mean ± SE (log arc seconds)	2.49 ± 0.21		2.71 ± 0.21	
Amblyopic eye's spherical equivalent (diopters)				
Mean ± SE	3.95 ± 0.67		4.05 ± 0.45	
Fellow eye's spherical equivalent (diopters)				
Mean ± SE	0.85 ± 0.39		0.77 ± 0.36	

logMAR = logarithm of the minimum angle of resolution; SE = standard error

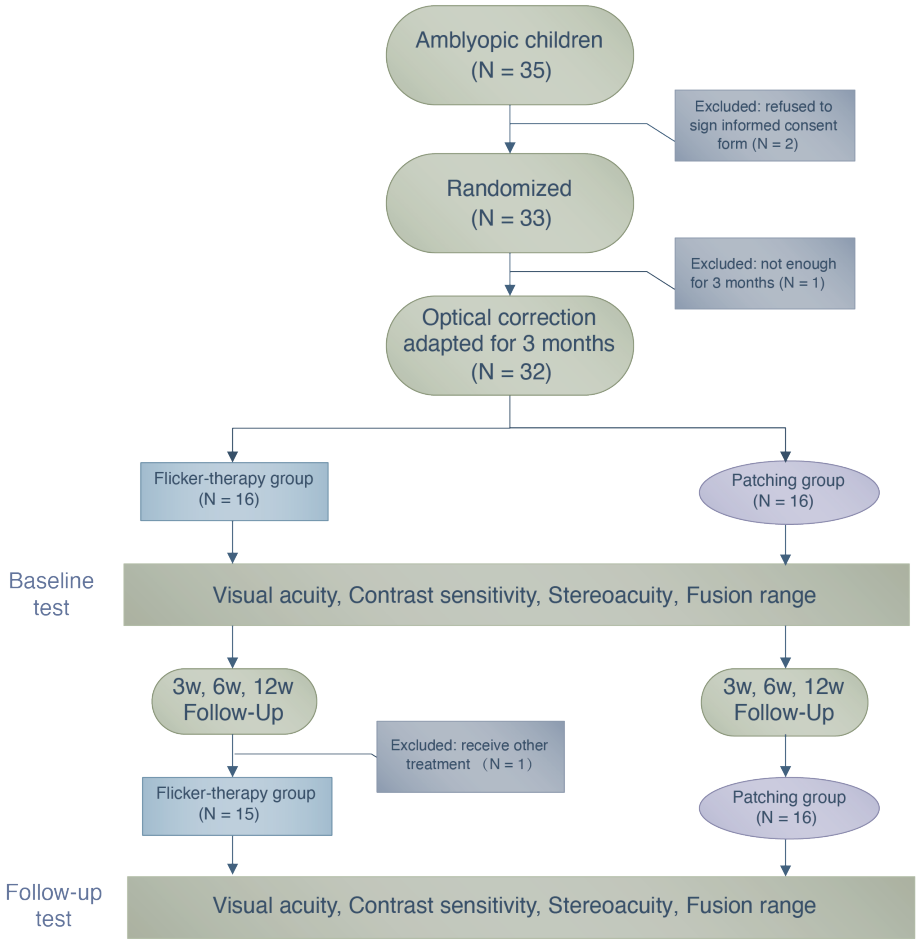

Figure 2: **A flow chart illustrating the treatment procedure and the number of patients who participated in this study.**

Visual acuity: Best-corrected visual acuity (BVCA) were measured separately for each eye using a national standard visual acuity chart according to the criteria for visual acuity test established by the National Health Commission of the People's Republic of China (Logarithmic Visual Acuity Chart, xk100-06, CHINA). The total score within each line from the Logarithmic Visual Acuity Chart was 0.1 log units. Since there were five letters per line, the correctly read letter was assigned a score of 0.02 log units. The patients were tested with one eye at 5 m from the chart and the non-tested eye occluded throughout the test.

Contrast sensitivity: A CSV-1000 grating chart (VectorVision® Inc., Greenville, OH 45331, USA) was used to measure the contrast sensitivity of the patients. This grating chart enabled us to test sensitivity for four spatial frequencies: 3, 6, 12 and 18 cycles per degree (cpd). This chart has been shown to be clinically reliable for tracking changes in visual performance of clinical patients [28]. We instructed the patients, who were equipped with optimal spectacle correction, to stand at a viewing distance of 40 cm in a dark room throughout the test.

Stereoacuity: Stereoacuity was assessed at a viewing distance of 40 cm under natural light with a Titmus Stereo test (Stereo Optical Co., Chicago, IL, USA). All patients wore polarizing glasses and, if necessary, additional optical corrections throughout testing. The Titmus Stereo test has been widely used in the clinical setting for measuring stereoacuity [29, 30]. If the patients were not able to perceive the largest disparity given by the Titmus Stereo test, we recorded that the stereoacuity of the patients was 'nil' and converted the non-numerical value into 4500 arc seconds for data analysis.

Fusion range: A good fusion range can stabilize binocular vision. As an index of binocular vision, we measured fusion range between both eyes in all patients using Synoptophore L-2510B/L-2510HB (Inami & Co., Ltd., Japan). The scales in the device for measuring fusion range are degrees and Prism Dioptres (Δ).

2.4 Data analysis

Statistical analyses were performed using R software [31]. Data were visualized using Python's matplotlib [32] and R's ggplot2 [33] packages. A Shapiro-Wilk test indicated that our dataset did not assume a normal distribution ($p < 0.05$) and failed the meet the requirement for the use of parametric procedures. Therefore, we performed nonparametric (rank-based) analysis of variance (ANOVA)-like computation of longitudinal data using the package 'nparLD' designed for R software [34]. It is a rank-based finite sample approximation based on the quantiles from F-distribution. Post-hoc tests were performed when there was a main effect in the within-subject factor time (baseline and follow-up measurement). P-values were corrected according to Holm-Bonferroni. We report ANOVA-type statistics [35]. When the interaction effect was significant, we interpreted our results based on the interaction effect rather than the main effect because some forms of interaction such as crossover interaction can show that a main effect exists even though they are not statistically significant.

As for the effect size, relative treatment effect (RTE) was computed using nparLD. RTE is a probability that a randomly chosen observation (i.e., subject) from a subgroup (i.e., one of the treatment groups) has larger values than that from the whole sample (combined distribution). For instance, an RTE of 0.6 for a subgroup indicates that there is an estimated probability of 60% for the value of a randomly drawn observation from the subgroup to be higher than that from the whole sample (i.e., mean distribution function). Therefore, an RTE has a value from 0 to 1 (as it is the case of probability). An RTE of 0.5 indicates no effect because it shows that the estimated probability for an observation within a subgroup to be higher than that in the mean distribution function is 50%. An RTE < 0.5 indicates that that there is a tendency for a randomly drawn subject within a subgroup to have a lower value than the counterpart from the whole sample. An RTE > 0.5 indicates that that there is a tendency for a randomly drawn subject within a subgroup to have a higher value than the counterpart from the whole sample. A 95% confidence interval of RTE that does not overlap with 0.5 indicates a significance ($p < 0.05$).

In most cases where contrast sensitivity across spatial frequency is measured in clinical studies, it is mainly of interest to see whether a novel treatment improves contrast sensitivity for a specific spatial frequency rather than the difference in contrast sensitivity across spatial frequencies. Therefore, in order to analyze our data, we performed the non-parametric ANOVA-like test in the form of the F1-LD-F1 design, which includes one between-subject factor (treatment group) and one within-subject factor (time), rather than with the F1-LD-F2 design, which includes one between-subject factor (treatment group) and two within-subject factors (time and spatial frequency). To perform the computation with a F1-LD-F1 model, we divided the entire dataset of contrast sensitivity into four groups as there were four different levels of spatial frequency (3, 6, 12, 18 cpd). By reorganizing our data, we were able to ignore the differences in contrast sensitivity between spatial frequencies and measure the true RTEs that pertain to the differences in contrast sensitivity for a specific spatial frequency across treatment duration (i.e., time) in the two treatment groups. In the rest of our data (visual acuity, stereopsis and fusion range), we also applied the F1-LD-F1 model (between-subject factor = treatment group, within-subject factor = time) to perform the nonparametric ANOVA-like test.

Table 2: **Relative treatment effects, their 95% confidence intervals (CI) and variances of visual acuity of the amblyopic eye, stereopsis and fusion range**

Visual function	Treatment group	Time point	Relative treatment effect	95% CI	Variance
Visual acuity	Flicker therapy	Baseline	**0.68***	**0.57-0.76**	0.0778
		3-week	0.56	0.45-0.65	0.0885
		6-week	0.52	0.42-0.62	0.0842
		12-week	0.50	0.41-0.60	0.0822
	Patching therapy	Baseline	**0.63***	**0.54-0.72**	0.0712
		3-week	0.45	0.36-0.54	0.0717
		6-week	**0.36***	**0.26-0.47**	0.0920
		12-week	**0.29***	**0.19-0.43**	0.114
Stereopsis	Flicker therapy	Baseline	0.50	0.38-0.62	0.125
		3-week	0.46	0.35-0.57	0.108
		6-week	0.49	0.38-0.60	0.104
		12-week	0.41	0.31-0.51	0.0893
	Patching therapy	Baseline	0.58	0.47-0.69	0.0987
		3-week	0.54	0.43-0.64	0.0909
		6-week	0.51	0.40-0.62	0.0983
		12-week	0.51	0.40-0.61	0.0890
Fusion range	Flicker therapy	Baseline	0.61	0.48-0.73	0.116
		3-week	0.36	0.24-0.51	0.133
		6-week	0.42	0.31-0.54	0.0970
		12-week	0.45	0.33-0.58	0.110
	Patching therapy	Baseline	0.62	0.49-0.73	0.0982
		3-week	0.50	0.38-0.63	0.104
		6-week	0.56	0.43-0.69	0.124
		12-week	0.46	0.32-0.62	0.161

*denote significance of the data (i.e., no overlap with RTE of 0.05 in the confidence interval). The significance is equivalent to $p < 0.05$.

3 Results

3.1 Baseline characteristics

We wanted to ensure that all patients were randomly assigned to the two treatment groups. So, by conducting statistical tests, we looked for possible differences between the Flickering and Patching groups based on the baseline visual performance. To do so, we used the nonparametric Wilcoxon-Mann-Whitney test. We found that the visual performance did not differ between the treatment groups (p-values ranged from 0.35 to 0.96) and concluded that the group assignment was random.

3.2 Compliance

We computed compliance by dividing the number of actual days where the patients had undergone the treatment at home from the number of total days for treatment. The mean compliance rate in the Flicker Group was $93.80 \pm 0.025\%$, whereas the mean compliance rate in the Patching Group was $93.97 \pm 0.021\%$. A nonparametric Wilcoxon-Mann-

Table 3: **Relative treatment effects, their 95% confidence intervals (CI) and variances of contrast sensitivity within the amblyopic eye**

Visual function	Treatment group	Spatial frequency	Time point	Relative treatment effect	95% CI	Variance
Contrast sensitivity	Flicker therapy	3 cpd	Baseline	0.38	0.27-0.50	0.105
			3-week	0.46	0.34-0.58	0.117
			6-week	0.58	0.47-0.68	0.0959
			12-week	0.56	0.44-0.66	0.106
	Patching therapy	3 cpd	Baseline	**0.31***	**0.22-0.44**	0.102
			3-week	0.55	0.44-0.65	0.0849
			6-week	0.57	0.44-0.68	0.127
			12-week	0.60	0.48-0.72	0.122
	Flicker therapy	6 cpd	Baseline	0.39	0.29-0.52	0.113
			3-week	0.52	0.39-0.64	0.134
			6-week	0.56	0.43-0.67	0.124
			12-week	0.60	0.47-0.71	0.116
	Patching therapy	6 cpd	Baseline	**0.36***	**0.27-0.46**	0.0762
			3-week	0.39	0.28-0.51	0.112
			6-week	0.56	0.45-0.67	0.102
			12-week	0.63	0.52-0.73	0.0954
	Flicker therapy	12 cpd	Baseline	0.38	0.27-0.50	0.111
			3-week	0.57	0.46-0.68	0.0980
			6-week	0.51	0.39-0.63	0.121
			12-week	0.62	0.50-0.72	0.0941
	Patching therapy	12 cpd	Baseline	**0.33***	**0.23-0.44**	0.0928
			3-week	0.49	0.38-0.60	0.105
			6-week	0.54	0.43-0.65	0.102
			12-week	0.57	0.45-0.68	0.110
	Flicker therapy	18 cpd	Baseline	0.40	0.28-0.55	0.157
			3-week	0.53	0.39-0.66	0.157
			6-week	0.53	0.41-0.64	0.106
			12-week	0.57	0.45-0.68	0.106
	Patching therapy	18 cpd	Baseline	0.38	0.28-0.50	0.0960
			3-week	0.52	0.42-0.61	0.0763
			6-week	0.50	0.40-0.61	0.0913
			12-week	0.56	0.44-0.68	0.117

*denote significance of the data (i.e., no overlap with RTE of 0.05 in the confidence interval). The significance is equivalent to $p < 0.05$.

Whitney test revealed no significant difference in compliance rate between the two treatment groups (W = 116, p = 0.89).

3.3 Comparison between the EFG therapy and the standard patching therapy

Our study attempts to clarify whether the EFG therapy (i.e., Flicker Group) could be a robust and reliable treatment that could perhaps replace the standard patching therapy (i.e., Patching Group). To upend the standard treatment, it would first have to show whether the patients who had undergone the EFG therapy would show improvements in monocular and binocular visual functions. Second, it would have to bring about a significantly larger improvement than that from the standard patching therapy. Detailed statistical analyses are outlined below and revolve around these two central issues.

3.3.1 Monocular visual functions

3.3.1.1 Visual acuity

To quantify whether the improvement in visual acuity from the Flicker group was significant relative to baseline and to see whether the difference in the magnitude of visual acuity improvement between the two treatment groups in the improvement was significant, we conducted a nonparametric ANOVA-like test. It revealed no main effect of group ($F_{(1.00,\infty)} = 2.09$, $p = 0.15$; RTEs = 0.56 and 0.44 for Flicker Group and Patching Group) and a significant effect of time ($F_{(2.29,\infty)} = 41.16$, $p < 0.001$; RTEs = 0.66, 0.50, 0.44, 0.40 for baseline, 3-week, 6-week and 12-week follow-up visits). Then, we performed post hoc tests for a significant main effect of within-subject factor time. We observed a significant difference between baseline and the subsequent follow-up visits (p's « 0.0001), and between 3-week and 12-week ($F_{(1.00,\infty)} = 12.97$, $p < 0.0001$). The test also revealed a significant interaction between the treatment group and time ($F_{(2.29,\infty)} = 4.39$, $p = 0.0091$; RTEs for Flicker Group = 0.68, 0.56, 0.52, 0.50 for baseline, 3-week, 6-week and 12-week follow-up visits; RTEs for Patching Group = 0.64, 0.45, 0.36, 0.29 for baseline, 3-week, 6-week and 12-week follow-up visits). This significant interaction effect is illustrated in Figure 4, which shows non-parallel time curves of logMAR visual acuity score in Flicker Group and Patching Group. The non-parallel plots in the function of time suggest that the time course of improvements in visual acuity differed between the two treatment groups. In fact, the improvement in visual acuity was larger in the Patching group than in the Flicker Group at certain, but not all, time points, such as at week 12.

Although no main effect of treatment group was observed, Patching Group's RTE of 0.44 was less than 0.50, suggesting that standard patching treatment resulted in lower score of logMAR visual acuity throughout the treatment period (and hence, more improvement) than that of a randomly drawn patient from our entire sample. On the other hand, Flicker Group's RTE of 0.56 exceeds 0.50, indicating that EFG treatment resulted in a higher score of logMAR visual acuity than that of a randomly drawn patient from our entire sample. This difference of RTE between the two treatment groups indicate the more notable efficacy of the standard patching treatment, indicated by the lower logMAR visual acuity scores. As for the effect of time, the RTEs at 6-week and 12-week follow-up visits (0.44 and 0.40, respectively) show that longer treatment period lowered the logMAR visual acuity score than the shorter treatment period of 3 weeks or no treatment (i.e., baseline). The RTEs for the interaction between Flicker Group and time are at around 0.5 or above, whereas those for the interaction between the Patching Group and time are progressively decreasing over time. From the RTEs, we can note that the patients in the Patching Group attained a larger improvement, and therefore a lower score in logMAR visual acuity as the treatment period increased, whereas the decrease of the RTEs in the Flicker Group over treatment period is less noticeable. This difference is illustrated in Figure 4, where the logMAR visual acuity of the Flicker Group is higher (and hence, worse) than that of the Patching Group at 12-week timepoint.

The confidence intervals of the RTEs can be informative (see Table 2). For instance, the confidence intervals of the Patching Group at 6- and 12-week measurement do not overlap with 0.5, indicating that a significant improvement in the logMAR visual acuity score in the Patching Group. On the other hand, the confidence intervals of the RTEs from

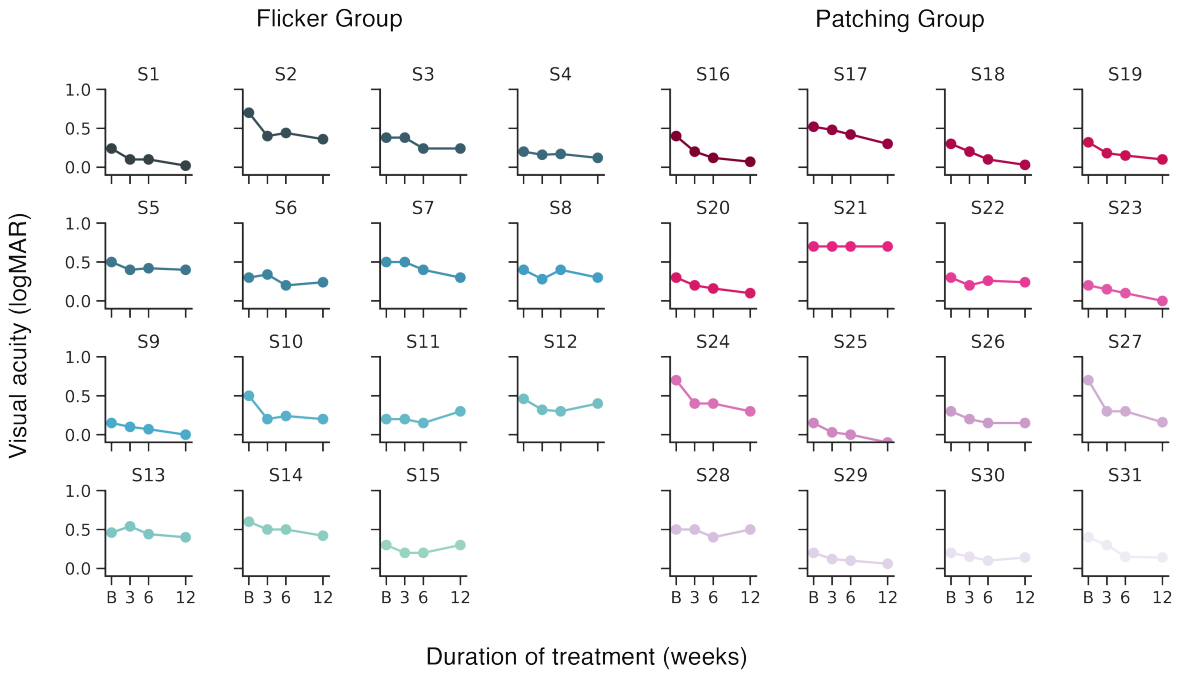

Figure 3: **Visual acuity of each patient's amblyopic eye at baseline and after beginning the treatment.** 'B' denotes baseline measurement. Y-axis represents logMAR visual acuity, x-axis represent the time of measurement (in weeks) throughout the study.

the Flicker Group overlap with 0.5 at all post-treatment timepoints, suggesting that the improvement is not as significant as the one in the Patching Group. However, the smaller improvement from the Flicker Group does not mean that the treatment was ineffective because the confidence interval of its baseline RTE does not overlap with 0.5. This indicates that the visual acuity score in baseline is significantly higher (i.e., worse visual acuity before the treatment) than the other time points.

3.3.1.2 Contrast sensitivity

Averaged data are shown in Figure 5. We were interested in determining whether contrast sensitivity of the amblyopic eye significantly improved as a result of the Flicker therapy relative to baseline, and to see whether the difference in improvement of contrast sensitivity differed significantly between the two treatment groups. To do so, we performed the non-parametric ANOVA-like test in the form of the F1-LD-F1 design, which includes one between-subject factor (treatment group) and one within-subject factor (time), rather than with the F1-LD-F2 design, which includes one between-subject factor (treatment group) and two within-subject factors (time and spatial frequency). In other words, instead of including spatial frequency as within-subject factor, we divided the datasets into four datasets to remove spatial frequency as within-subject factor. The rationales are described in detail in Methods. Our findings differ across spatial frequency. At 3 cpd, non-parametric ANOVA-like computation revealed no main effect of group ($F_{(1.00,\infty)} = 0.042$, $p = 0.84$; RTEs = 0.49 and 0.51 for the Flicker Group and the Patching Group) and a significant effect of time ($F_{(2.39,\infty)} = 8.84$, $p < 0.001$; RTEs = 0.35, 0.50, 0.57, 0.58 for baseline, 3-week, 6-week and 12-week follow-up visits). Then, we performed post-hoc tests for a significant main effect of within-subject factor time. We observed a significant difference between baseline and the subsequent follow-up visits (p's < 0.05), and between 3-week and

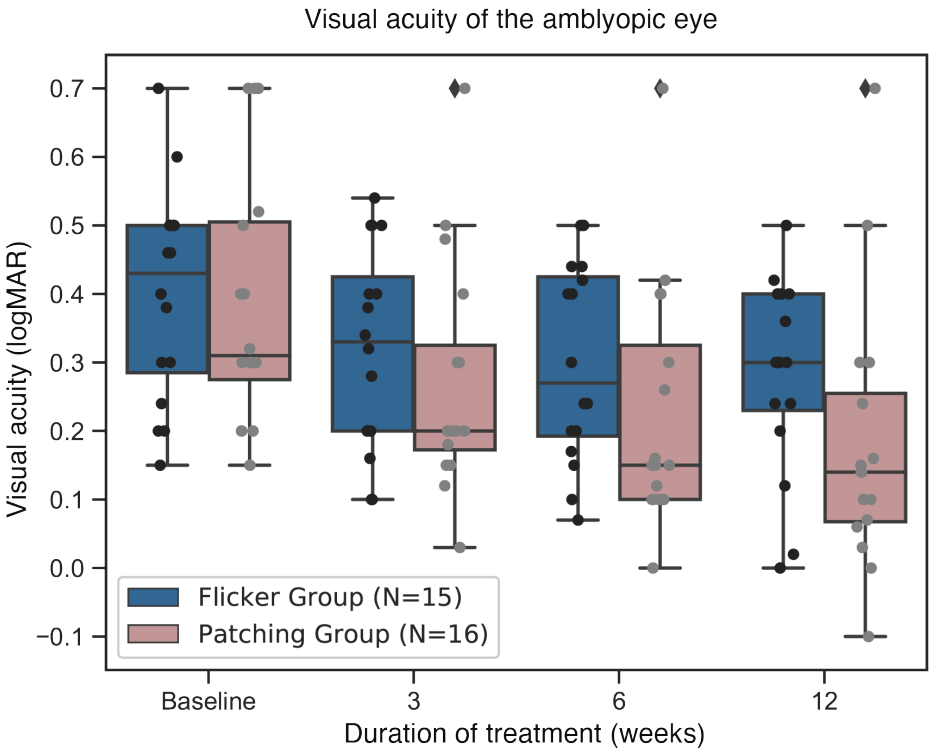

Figure 4: **Boxplots of logMAR visual acuity in Flicker and Patching treatment groups.** Blue represents the Flicker Group, pink the Patching Group. Individual data point is represented by a dot. The black solid line within each box represents the median. The box represents the interquartile range (IQR) of the data (25th to the 75th percentile). The whisker represents 1.5 × IQR either above the third quartile or below the first quartile. Diamonds represent outliers.

12-week ($F_{(1.00,\infty)} = 12.97$, $p = 0.0013$). The test did not reveal a significant interaction between the treatment group and time ($F_{(2.39,\infty)} = 0.89$, $p = 0.43$; RTEs for Flicker Group = 0.38, 0.46, 0.58, 0.56 for baseline, 3-week, 6-week and 12-week follow-up visits; RTEs for Patching Group = 0.31, 0.55, 0.57, 0.60 for baseline, 3-week, 6-week and 12-week follow-up visits). As for the effect of time, the RTEs at 6-week and 12-week follow-up visits (0.57 and 0.58, respectively) show that longer treatment period improved contrast sensitivity than the shorter treatment period of 3 weeks or no treatment (i.e., baseline). Conversely, the RTE of 0.35 at baseline indicates that contrast sensitivity is relatively low compared to that after beginning the treatment. The confidence intervals of the RTEs enable a more detailed understanding (see Table 3). For instance, the confidence interval of RTE at baseline in Flicker group overlaps with 0.5, whereas the one in the Patching group does not. Therefore, it seems that baseline contrast sensitivity is significantly lower in the Patching Group than the combined dataset of 3 cpd. This suggests that, at 3 cpd, patients' contrast sensitivities significantly improved as a result of the patching therapy, but not the flicker therapy.

At 6 cpd, non-parametric ANOVA-like computation revealed no main effect of group ($F_{(1.00,\infty)} = 0.16$, $p = 0.69$; RTEs = 0.52 and 0.49 for the Flicker Group and the Patching Group) and a significant effect of time ($F_{(3.00,\infty)} = 20.43$, $p < 0.001$; RTEs = 0.38, 0.45, 0.56, 0.62 for baseline, 3-week, 6-week and 12-week follow-up visits). Then, we performed post-hoc tests for a significant main effect of within-subject factor time. We observed a significant difference between baseline and 6-week ($F_{(1.00,\infty)} = 12.02$, $p = 0.0021$), between baseline and 12-week ($F(_{1.00,\infty)} = 17.83$, $p < 0.0001$) and between 3-week and 12-week ($F_{(1.00,\infty)} = 7.81$, $p = 0.021$). The test did not reveal a significant interaction between the treatment group and time ($F_{(2.74,\infty)} = 0.95$, $p = 0.41$; RTEs for Flicker Group = 0.39, 0.51, 0.56, 0.60 for baseline, 3-week, 6-week and 12-week follow-up visits; RTEs for Patching Group = 0.36, 0.39, 0.56, 0.63 for baseline, 3-week,

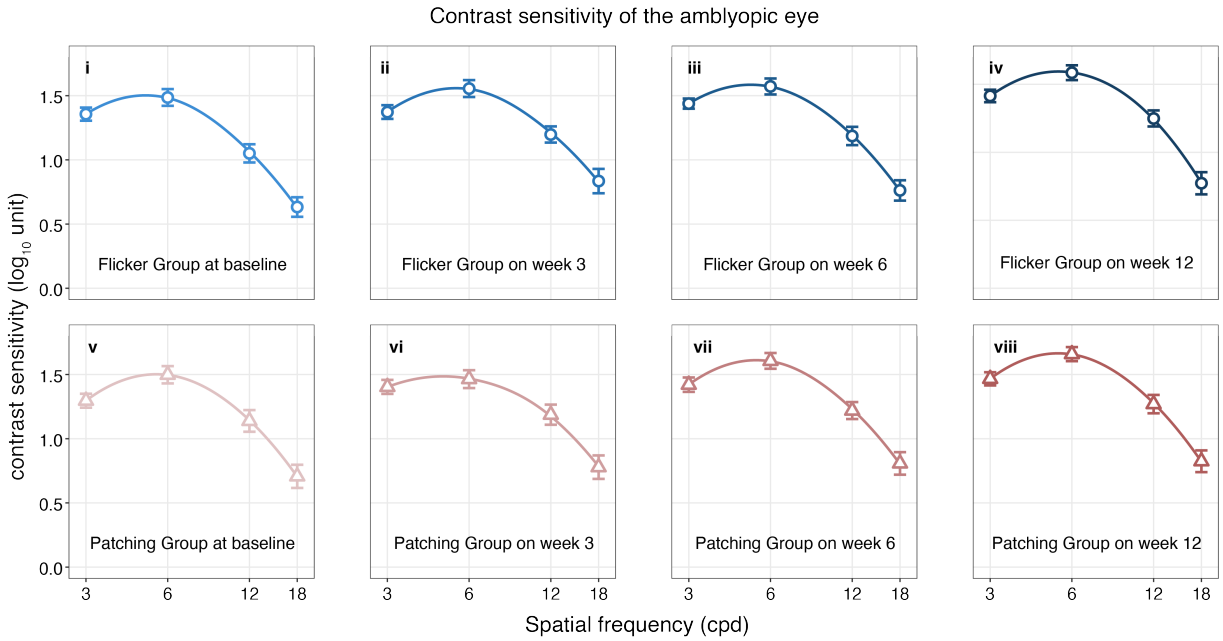

Figure 5: **Means of contrast sensitivity for each treatment group at all time points of the treatment procedure.** Blue plots represent the Flicker Group, pink plots the Patching Group. The points of circle represent the mean contrast sensitivity across patients in the Flicker Group, and the points of triangle the mean contrast sensitivity across patients in the Patching Group. The increased darkness for each color represents the increased period of treatment.

6-week and 12-week follow-up visits). As for the effect of time, the RTEs at 6-week and 12-week follow-up visits (0.56 and 0.62, respectively) show that longer treatment period improved contrast sensitivity than the shorter treatment period of 3 weeks or no treatment (i.e., baseline). Conversely, the RTE of 0.38 at baseline indicates that contrast sensitivity is relatively low compared to that in the follow-up visits. Moreover, the confidence interval of RTE at baseline in Flicker group overlaps with 0.5, whereas the one in Patching group does not (see Table 3). Therefore, it seems that baseline contrast sensitivity is significantly lower in Patching Group than the combined dataset of 6 cpd. This suggests that, at 6 cpd, patients' contrast sensitivities significantly improved as a result of the patching therapy, but not the flicker therapy.

At 12 cpd, non-parametric ANOVA-like computation revealed no main effect of group ($F_{(1.00,\infty)} = 0.16$, $p = 0.62$; RTEs = 0.52 and 0.48 for the Flicker Group and the Patching Group) and a significant effect of time ($F_{(2.36,\infty)} = 7.95$, $p < 0.001$; RTEs = 0.35, 0.53, 0.52, 0.59 for baseline, 3-week, 6-week and 12-week follow-up visits). Then, we performed post-hoc tests for a significant main effect of within-subject factor time. We observed a significant difference between baseline and all subsequent timepoints (p's < 0.05). The test did not reveal a significant interaction between the treatment group and time ($F(2.36,\infty) = 0.45$, $p = 0.67$; RTEs for Flicker Group = 0.38, 0.57, 0.51, 0.62 for baseline, 3-week, 6-week and 12-week follow-up visits; RTEs for Patching Group = 0.33, 0.49, 0.54, 0.57 for baseline, 3-week, 6-week and 12-week follow-up visits). As for the effect of time, the RTE at 12-week follow-up visit (0.59) show that longer treatment period improved contrast sensitivity than the shorter treatment periods of 3 and 6 weeks, as well as no treatment (i.e., baseline). Conversely, the RTE of 0.35 at baseline indicates that contrast sensitivity is relatively low compared to that in the follow-up visits. In addition, the confidence interval of RTE at baseline in the Flicker group overlaps with 0.5, whereas the one in the Patching group does not (see Table 3). Therefore, it seems that baseline contrast sensitivity is significantly lower in the Patching Group than the combined dataset of 12 cpd. This suggests that contrast sensitivity significantly improved as a result of the patching therapy, but not the flicker therapy.

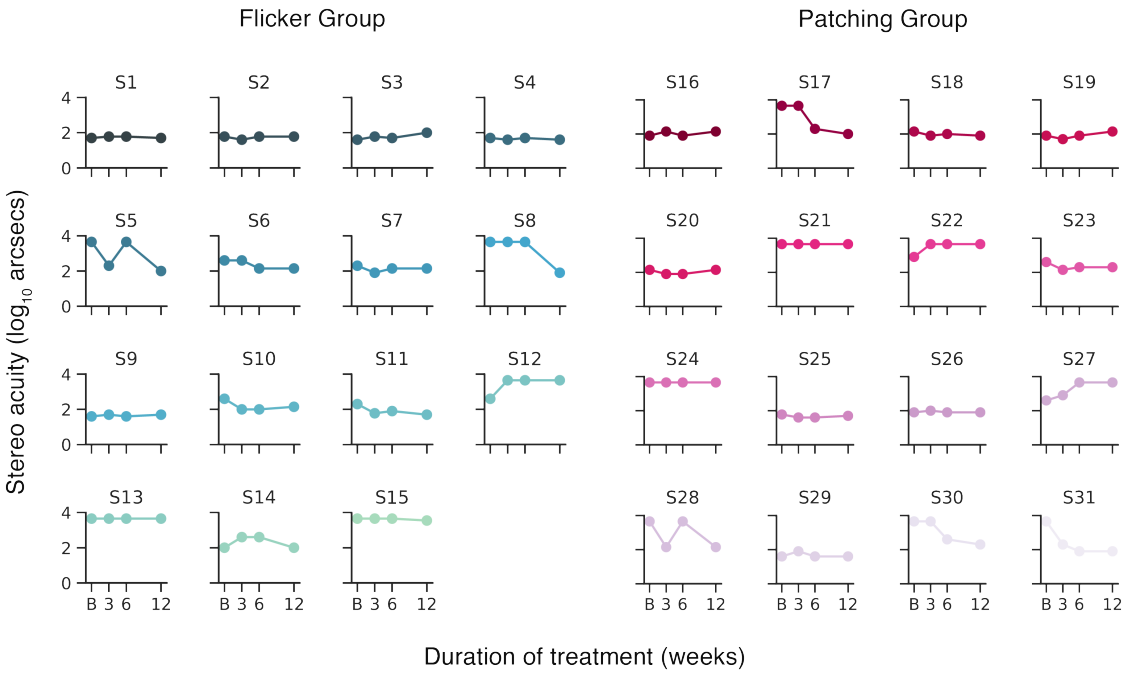

Figure 6: **Individual patient's stereoacuity at baseline and after beginning the treatment.** 'B' denotes baseline measurement. Y-axis represents stereo threshold that has been log transformed (log10 of the original stereo threshold), x-axis represent the time of measurement throughout the study.

At 18 cpd, nonparametric ANOVA-like computation revealed no main effect of group ($F_{(1.00,\infty)}$= 0.054, p = 0.82; RTEs = 0.51 and 0.49 for the Flicker Group and the Patching Group) and a significant effect of time ($F_{(2.44,\infty)}$ = 3.42, p = 0.024; RTEs = 0.39, 0.52, 0.51, 0.57 for baseline, 3-week, 6-week and 12-week follow-up visits). Then, we performed post hoc tests for a significant main effect of within-subject factor time. We observed a significant difference between baseline and week 12 measurements ($F_{(1.00,\infty)}$ = 6.44, p = 0.011). The test did not reveal a significant interaction between the treatment group and time ($F_{(2.43,\infty)}$ = 0.014, p = 0.99; RTEs for the Flicker Group = 0.40, 0.53, 0.53, 0.57 for baseline, 3-week, 6-week and 12-week follow-up visits; RTEs for the Patching Group = 0.38, 0.52, 0.50, 0.56 for baseline, 3-week, 6-week and 12-week follow-up visits). As for the effect of time, the RTE at 12-week follow-up visit (0.57) show that longer treatment period improved contrast sensitivity than the shorter treatment periods of 3 and 6 weeks, as well as no treatment (i.e., baseline). Conversely, the RTE of 0.39 at baseline indicates that contrast sensitivity is relatively low compared to that in the follow-up visits. Furthermore, the confidence intervals of RTE at baseline in Flicker group and Patching Group overlap with 0.5 (see Table 3). Therefore, it seems that baseline contrast sensitivity is not significantly lower in both treatment groups than a randomly drawn observation in the combined dataset of 18 cpd. This suggests although contrast sensitivity significantly improved at other spatial frequencies because of both therapies, contrast sensitivity at 18 cpd did not improve significantly.

3.3.2 Binocular visual functions

3.3.2.1 Stereopsis

We were interested in determining whether stereopsis significantly improved as a result of the Flicker therapy relative to baseline, and to see whether the difference in improvement of stereopsis differed significantly between the two treatment

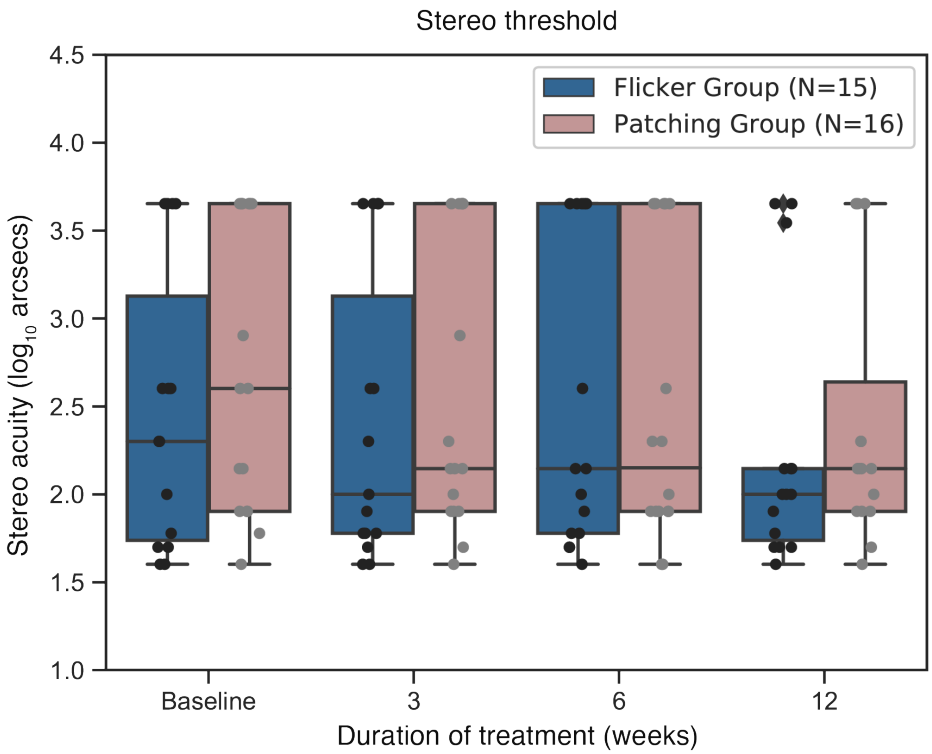

Figure 7: **Boxplots of stereo thresholds in Flicker and Patching treatment groups.** Blue represents the Flicker Group, pink the Patching Group. Individual data point is represented by a dot. The black solid line within each box represents the median. The box represents the interquartile range (IQR) of the data (25th to the 75th percentile). The whisker represents 1.5 × IQR either above the third quartile or below the first quartile. Diamonds represent outliers.

groups. Nonparametric ANOVA-like computation revealed no main effect of group ($F_{(1.00,\infty)} = 0.52$, $p = 0.47$; RTEs = 0.46 and 0.53 for the Flicker Group and the Patching Group) and no significant effect of time ($F_{(2.46,\infty)} = 2.18$, $p = 0.10$; RTEs = 0.54, 0.50, 0.50, 0.46 for baseline, 3-week, 6-week and 12-week follow-up visits). Also, it revealed a no interaction effect between the treatment group and time ($F_{(2.46,\infty)} = 0.56$, $p = 0.61$; RTEs for the Flicker Group = 0.50, 0.46, 0.49, 0.41 for baseline, 3-week, 6-week and 12-week follow-up visits; RTEs for the Patching Group = 0.58, 0.54, 0.51, 0.51 for baseline, 3-week, 6-week and 12-week follow-up visits). Confidence intervals of the RTEs are provided in Table 2 but they do not reveal any significant improvements in stereopsis throughout the treatment in both groups.

3.3.2.2 Fusion range

We were interested in determining whether fusion range significantly improved as a result of the Flicker therapy relative to baseline, and to see whether the difference in improvement of fusion range differed significantly between the two treatment groups. Nonparametric ANOVA-like computation revealed no main effect of group ($F(1.00,\infty) = 0.98$, $p = 0.32$; RTEs = 0.46 and 0.53 for the Flicker Group and the Patching Group) and a significant effect of time ($F_{(2.60,\infty)} = 3.27$, $p = 0.026$; RTEs = 0.62, 0.43, 0.49 0.46 for baseline, 3-week, 6-week and 12-week follow-up visits). Then, we performed post hoc tests for a significant main effect of within-subject factor time. We observed a significant difference between baseline and 3-week ($F_{(1.00,\infty)} = 7.99$, $p = 0.019$), and between baseline and 12-week ($F_{(1.00,\infty)} = 7.65$, $p = 0.022$). Also, it revealed a no interaction effect between the treatment group and time ($F_{(2.60,\infty)} = 0.77$, $p = 0.50$; RTEs for the Flicker Group = 0.61, 0.36, 0.42, 0.45 for baseline, 3-week, 6-week and 12-week follow-up visits; RTEs for

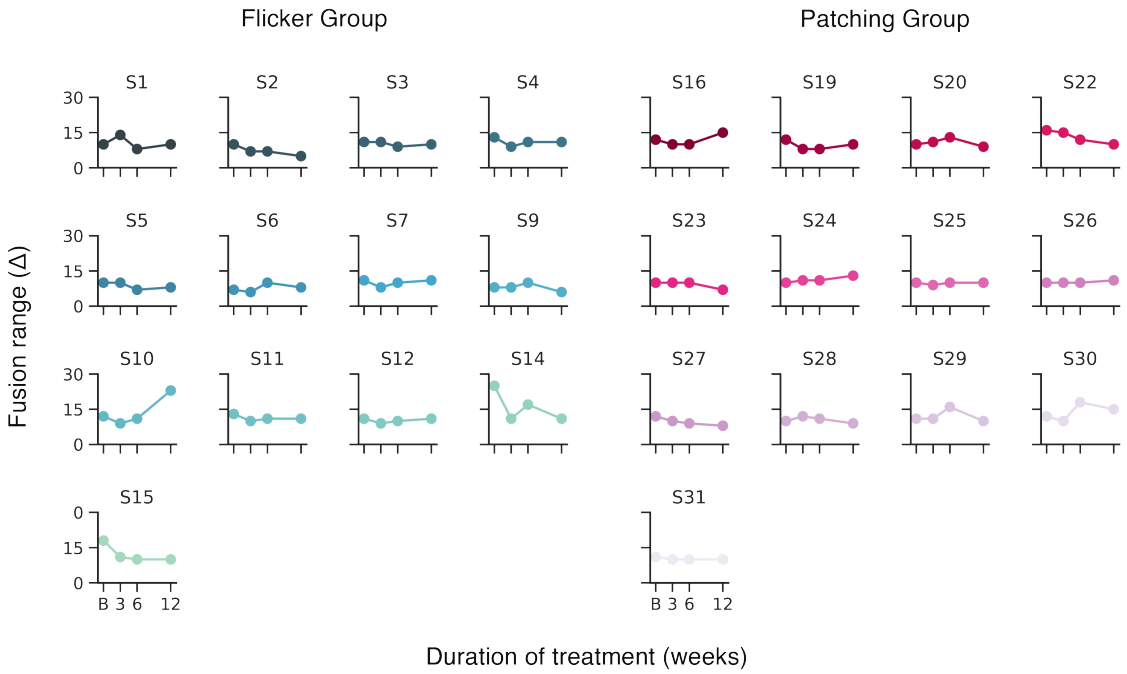

Figure 8: **Individual patient's fusion range at baseline and after beginning the treatment.** 'B' denotes baseline measurement. Y-axis represents fusion range, x-axis represent the time of measurement throughout the study.

the Patching Group = 0.62, 0.50, 0.56, 0.46 for baseline, 3-week, 6-week and 12-week follow-up visits). Confidence intervals of the RTEs are provided in Table 2 but they do not reveal any significant improvements in fusion range throughout the treatment in both groups.

4 Discussion

According to the literature within our grasp, this study marks the first randomized controlled trial that compares the efficacy of Eyetronix Flicker Glass and adhesive occlusion patches in children with anisometropic amblyopia. Our results show that the EFG therapy improves the visual acuity of the amblyopic eye significantly but does not so for the contrast sensitivity of the amblyopic eye. Moreover, standard patching therapy using adhesive occlusion patches seems to be more effective in improving visual acuity and contrast sensitivity (at 3, 6 and 12 cpd) of the amblyopic eye than the EFG therapy. On the other hand, both therapies did not improve fusion range and stereopsis after 12 weeks of daily treatment. The compliance rate between the two therapy groups was similar (Flicker Group: 93.80 $\pm$ 0.025%, Patching Group: 93.97 $\pm$ 0.021%). Our findings demonstrate that although the EFG therapy is a binocular treatment that deprives visual input through both eyes, it does not improve binocular functions. Also, EFG therapy fails to bring about a larger benefit in monocular visual functions than the standard patching therapy in the amblyopic population.

Improvements in monocular visual functions are larger from the standard patching therapy than the EFG therapy. The different deprivation lengths of the normal eye between the two treatments might account for the difference. To illustrate, deprivation of the normal eye for two hours (from two hours of the standard therapy) might have induced a larger improvement than the deprivation of the fellow eye for mere 30 minutes (from one hour of the EFG therapy). We had designed this study according to the recommendation by Eyetronix® that says 1-2 hours of daily treatment is

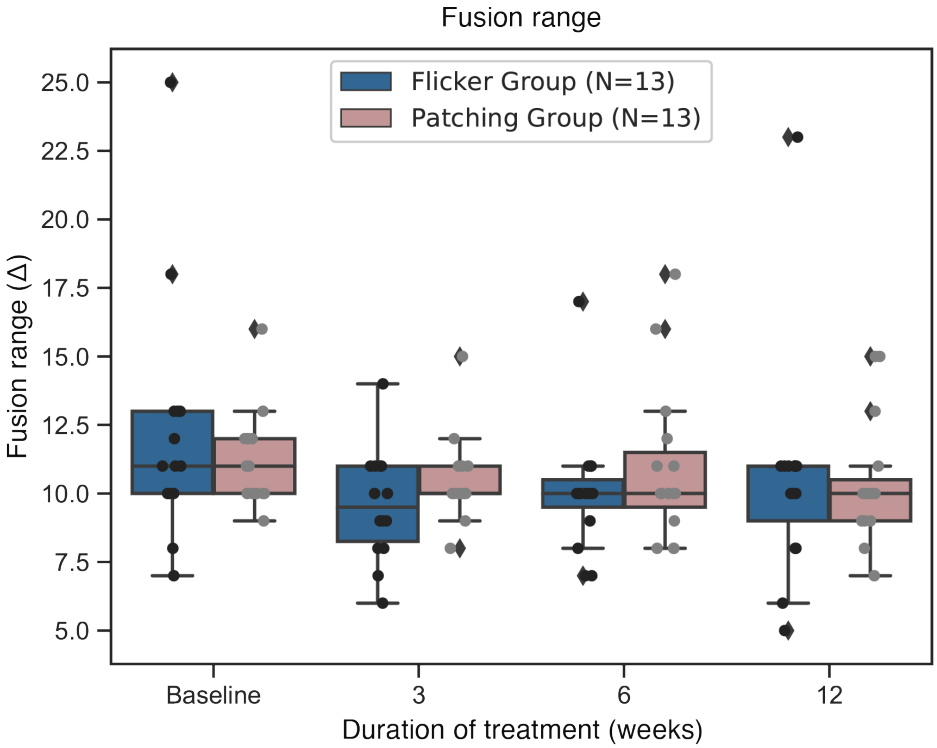

Figure 9: **Boxplots of fusion range in Flicker and Patching treatment groups.** Blue represents the Flicker Group, pink the Patching Group. Individual data point is represented by a dot. The black solid line within each box represents the median. The box represents the interquartile range (IQR) of the data (25th to the 75th percentile). The whisker represents 1.5 $\times$ IQR either above the third quartile or below the first quartile. Diamonds represent outliers.

adequate for visual improvements (http://eyetronix.com). However, we were not able to find studies that investigate the effect of EFG therapy duration in amblyopic patients.

Both therapies did not improve binocular visual functions, such as stereopsis and fusion range. EFG therapy was introduced on the premise that alternating visual flicker could reduce suppression and improve binocular vision [25]. However, our findings show no improvement in binocular visual functions from the EFG therapy. Our study therefore reproves the notion that binocular treatment strategy, which manipulates visual input in both eyes, does not guarantee binocular benefits for amblyopic patients. However, our findings are limited in scope because we only investigated binocular functions that reflect excitatory connections that drive binocular summation, such as stereopsis and fusion range. Inhibitory interactions between the amblyopic eye and normal eyes also shape binocular functions, such as interocular suppression. Chen et al. (2019) recently demonstrated that standard patching therapy does not ameliorate interocular suppression in the amblyopic population [11]. Indeed, as standard patching therapy is monocular treatment since it only manipulates visual input through the normal eye, this result is not surprising. One could speculate that only monocular visual functions would be influenced as a result of undergoing monocular treatment such as the standard patching therapy. A future study should also investigate whether the EFG therapy, which involves an alternate deprivation of the normal and fellow eyes, reduces interocular suppressions in amblyopic patients.

Eyetronix® introduced the EFG on the premise that manipulation of visual input through both eyes could redress the imbalanced binocular functions in amblyopia. A previous study by Vera-Diaz et al reports improvements in both visual acuity for the amblyopic eye and stereopsis after 12 weeks of EFG therapy in 23 children (aged 5-17 years) [27]. Although our results agree in that visual acuity of the amblyopic eye was improved after EFG therapy (i.e., monocular

function), they disagree in that stereopsis was not improved (i.e., binocular function). This can be due to multiple reasons. 1) Occlusion period was shorter in our study. The two previous studies employed 1-2 hours of occlusion as opposed to a fixed number of hours like in our study. 2) The age distribution of the patients seems to be notably different between our study and the previous study. For instance, all fifteen patients (in Flicker Group) from our study were aged four to nine years old. On the other hand, twelve out of twenty-three patients in the study of Vera-Diaz et al. (2016) were above ten years old [27]. However, despite the different age distributions in the patients across the studies, all of them had anisometropic amblyopia. More research groups should investigate the efficacy of the EFG therapy in a wide clinical population of amblyopia to improve the statistical power and replicability. Also, the effect of deprivation duration during the EFG therapy should be studied in a future study. Moreover, shorter and longer treatment durations of the EFG therapy could also be explored in future studies with the hope to find the optimal duration of the therapy.

5 Limitations

Our study has several limitations. First, since all of our patients were anisometropic, the implications of our findings might not be generalizable to other types of amblyopia, such as strabismic and deprivation amblyopia. Second, we did not equally deprive the normal eye for each treatment group. The patients in the Patching Group underwent 2-hour deprivation of the fellow eye and those in the Flicker Group 30-minute deprivation of the fellow eye (in a 1-hour treatment). Third, we did not perform follow-up assessments of the patients after the end of the treatment. Therefore, we can conclude that standard patching therapy is superior to the EFG therapy only during but not after the treatment. Whether the monocular benefit from the EFG therapy would last remains to be investigated.

6 Conclusion

In children aged 4-13 years old with moderate or several unilateral anisometropic amblyopia, the standard patching therapy (i.e., patching the fellow eye for two hours/day) improved visual acuity and contrast sensitivity of the amblyopic eye more so than the EFG therapy (i.e., alternative flicker for one hour/day). Both therapies did not improve fusion range and stereopsis after 12 weeks of daily treatment. Therefore, our study suggests that although the EFG therapy incorporates a binocular treatment strategy, whereby visual input in both eyes is manipulated, it does not ensure binocular benefits in amblyopic patients. Coupled with a smaller improvement in monocular visual functions that was observed from the EFG therapy, whether EFG therapy could replace the standard patching therapy seems to be questionable.

7 Acknowledgment

This research was supported by National Natural Science Foundation of China (Grant Nos. 31970975 and 81500754) and the Wenzhou Medical University grant (QTJ16005) to JWZ and Canadian Institute of Health Research's graduate award to SHM.

References

[1] Jonathan M Holmes and Michael P Clarke. Amblyopia. *The Lancet*, 367(9519):1343–1351, 2006.

[2] Robert F Hess and Benjamin Thompson. Amblyopia and the binocular approach to its therapy. *Vision research*, 114:4–16, 2015.

[3] Dennis M Levi, David C Knill, and Daphne Bavelier. Stereopsis and amblyopia: A mini-review. *Vision research*, 114:17–30, 2015.

[4] C De Buffon. Disseration sur la cause du strabisme ou des yeux louches. *Me. Acad Roy Sci*, 1743, 1743.

[5] Mitchell M Scheiman, Richard W Hertle, Roy W Beck, Allison R Edwards, Eileen Birch, Susan A Cotter, Earl R Crouch Jr, Oscar A Cruz, Bradley V Davitt, Sean Donahue, et al. Randomized trial of treatment of amblyopia in children aged 7 to 17 years. *Archives of ophthalmology (Chicago, Ill.: 1960)*, 123(4):437–447, 2005.

[6] Maria Fronius, Licia Cirina, Hanns Ackermann, Thomas Kohnen, and Corinna M Diehl. Efficiency of electronically monitored amblyopia treatment between 5 and 16 years of age: new insight into declining susceptibility of the visual system. *Vision research*, 103:11–19, 2014.

[7] Catherine E Stewart, Merrick J Moseley, David A Stephens, and Alistair R Fielder. Treatment dose-response in amblyopia therapy: the monitored occlusion treatment of amblyopia study (motas). *Investigative ophthalmology & visual science*, 45(9):3048–3054, 2004.

[8] Michael P Wallace, Catherine E Stewart, Merrick J Moseley, David A Stephens, and Alistair R Fielder. Compliance with occlusion therapy for childhood amblyopia. *Investigative ophthalmology & visual science*, 54(9):6158–6166, 2013.

[9] A Searle, P Norman, R Harrad, and K Vedhara. Psychosocial and clinical determinants of compliance with occlusion therapy for amblyopic children. *Eye*, 16(2):150–155, 2002.

[10] Eileen E Birch. Amblyopia and binocular vision. *Progress in retinal and eye research*, 33:67–84, 2013.

[11] Yiya Chen, Zhifen He, Yu Mao, Hao Chen, Jiawei Zhou, and Robert F Hess. Patching and suppression in amblyopia: one mechanism or two? *Frontiers in Neuroscience*, 13, 2019.

[12] Eileen E Birch, Yolanda S Castañeda, Christina S Cheng-Patel, Sarah E Morale, Krista R Kelly, Cynthia L Beauchamp, and Ann Webber. Self-perception of school-aged children with amblyopia and its association with reading speed and motor skills. *JAMA ophthalmology*, 137(2):167–174, 2019.

[13] Stephen B Kaye, Sean I Chen, Gary Price, Lesley C Kaye, Carmel Noonan, Ajay Tripathi, Pammal Ashwin, Nolan Cota, David Clark, and Jeremy Butcher. Combined optical and atropine penalization for the treatment of strabismic and anisometropic amblyopia. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, 6(5):289–293, 2002.

[14] Omry BenEzra, Rafi Herzog, Evelyne Cohen, Ilana Karshai, and David BenEzra. Liquid crystal glasses: feasibility and safety of a new modality for treating amblyopia. *Archives of ophthalmology*, 125(4):580–581, 2007.

[15] Omry BenEzra, Abraham Spierer, and Oriel Spierer. Liquid crystals—transformer that meets the eye. *Journal of pediatric ophthalmology and strabismus*, 52(5):266–267, 2015.

[16] Ibrahim Erbağcı, Seydi Okumuş, Veysi Öner, Erol Coşkun, Oğuz Çelik, and Burak Ören. Using liquid crystal glasses to treat amblyopia in children. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, 19(3):257–259, 2015.

[17] Abraham Spierer, Judith Raz, Omry BenEzra, Rafi Herzog, Evelyne Cohen, Ilana Karshai, and David BenEzra. Treating amblyopia with liquid crystal glasses: a pilot study. *Investigative ophthalmology & visual science*, 51(7):3395–3398, 2010.

[18] Daniel H Baker, Tim S Meese, and Robert F Hess. Contrast masking in strabismic amblyopia: attenuation, noise, interocular suppression and binocular summation. *Vision research*, 48(15):1625–1640, 2008.

[19] Jian Ding, Stanley A Klein, and Dennis M Levi. Binocular combination in abnormal binocular vision. *Journal of vision*, 13(2):14–14, 2013.

[20] Karen Holopigian, Randolph Blake, and Mark J Greenwald. Selective losses in binocular vision in anisometropic amblyopes. *Vision research*, 26(4):621–630, 1986.

[21] Chang-Bing Huang, Jiawei Zhou, Zhong-Lin Lu, and Yifeng Zhou. Deficient binocular combination reveals mechanisms of anisometropic amblyopia: Signal attenuation and interocular inhibition. *Journal of vision*, 11(6):4–4, 2011.

[22] MiYoung Kwon, Emily Wiecek, Steven C Dakin, and Peter J Bex. Spatial-frequency dependent binocular imbalance in amblyopia. *Scientific reports*, 5:17181, 2015.

[23] Jiawei Zhou, Alexandre Reynaud, Zhimo Yao, Rong Liu, Lixia Feng, Yifeng Zhou, and Robert F Hess. Amblyopic suppression: Passive attenuation, enhanced dichoptic masking by the fellow eye or reduced dichoptic masking by the amblyopic eye? *Investigative ophthalmology & visual science*, 59(10):4190–4197, 2018.

[24] Pi-Chun Huang, Daniel H Baker, and Robert F Hess. Interocular suppression in normal and amblyopic vision: spatio-temporal properties. *Journal of vision*, 12(11):29–29, 2012.

[25] Clifton Schor, Mark Terrell, and Donald Peterson. Contour interaction and temporal masking in strabismus and amblyopia. *American journal of optometry and physiological optics*, 53(5):217–223, 1976.

[26] Jingyun Wang, Daniel E Neely, Jay Galli, Joshua Schliesser, April Graves, Tina G Damarjian, Jessica Kovarik, James Bowsher, Heather A Smith, Dana Donaldson, et al. A pilot randomized clinical trial of intermittent occlusion therapy liquid crystal glasses versus traditional patching for treatment of moderate unilateral amblyopia. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, 20(4):326–331, 2016.

[27] Moore B. Hussey E. Srinivasan G. Johnson C. Vera-Diaz, F. A. A flicker therapy for the treatment of amblyopia. *Vision Development and Rehabilitation*, 2(2):105–114, 2016.

[28] Glenn N Pomerance and David W Evans. Test-retest reliability of the csv-1000 contrast test and its relationship to glaucoma therapy. *Investigative ophthalmology & visual science*, 35(9):3357–3361, 1994.

[29] Elise B Ciner, Gui-shuang Ying, Marjean Taylor Kulp, Maureen G Maguire, Graham E Quinn, Deborah Orel-Bixler, Lynn A Cyert, Bruce Moore, and Jiayan Huang. Stereoacuity of preschool children with and without vision disorders. *Optometry and vision science: official publication of the American Academy of Optometry*, 91(3):351, 2014.

[30] David K Wallace, Elizabeth L Lazar, Michele Melia, Eileen E Birch, Jonathan M Holmes, Kristine B Hopkins, Raymond T Kraker, Marjean T Kulp, Yi Pang, Michael X Repka, et al. Stereoacuity in children with anisometropic amblyopia. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, 15(5):455–461, 2011.

[31] J Allaire. Rstudio: integrated development environment for r. *Boston, MA*, 770:394, 2012.

[32] John D Hunter. Matplotlib: A 2d graphics environment. *Computing in science & engineering*, 9(3):90–95, 2007.

[33] Hadley Wickham. *ggplot2: elegant graphics for data analysis*. springer, 2016.

[34] Kimihiro Noguchi, Yulia R Gel, Edgar Brunner, and Frank Konietschke. nparld: an r software package for the nonparametric analysis of longitudinal data in factorial experiments. *Journal of Statistical Software*, 50(12), 2012.

[35] Edgar Brunner, Frank Konietschke, Markus Pauly, and Madan L Puri. Rank-based procedures in factorial designs: hypotheses about non-parametric treatment effects series b statistical methodology. 2017.