

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

Recovering stereo vision by squashing virtual bugs in a virtual reality environment

### Permalink

<https://escholarship.org/uc/item/6hb7321d>

### Journal

Philosophical Transactions of the Royal Society B Biological Sciences, 371(1697)

### ISSN

0962-8436

### Authors

Vedamurthy, Indu  
Knill, David C  
Huang, Samuel J  
[et al.](#)

### Publication Date

2016-06-19

### DOI

10.1098/rstb.2015.0264

Peer reviewed

# PHILOSOPHICAL TRANSACTIONS B

## Recovering stereo vision by squashing virtual bugs in a virtual reality environment.

Journal:	<i>Philosophical Transactions B</i>
Manuscript ID	RSTB-2015-0264.R2
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Vedamurthy, Indu; University of Rochester, Brain & Cognitive Sciences Knill, David; University of Rochester, Brain & Cognitive Sciences Huang, Sam; University of Rochester, Brain & Cognitive Sciences Yung, Amanda; University of Rochester, Brain & Cognitive Sciences Ding, Jian; University of California, Berkeley, Optometry Kwon, Oh-Sang; Ulsan National Institute of Science and Technology, School of Design & Human Engineering Bavelier, Daphne; University of Geneva, Faculty of Psychology and Education Sciences; University of Rochester, Brain & Cognitive Sciences Levi, Dennis; University of California, Berkeley, Optometry;
Issue Code: Click <a href="http://rstb.royalsocietypublishing.org/site/misc/issue-codes.xhtml">http://rstb.royalsocietypublishing.org/site/misc/issue-codes.xhtml</a> target=_new>here</a> to find the code for your issue.:	3DVIS
Subject:	Neuroscience < BIOLOGY
Keywords:	Stereopsis, Strabismus, Amblyopia, Virtual Reality, Perceptual learning, stereoblindness

SCHOLARONE™  
Manuscripts

# Recovering stereo vision by squashing virtual bugs in a virtual reality environment

Indu Vedamurthy<sup>1†</sup>, David C. Knill<sup>1</sup>, Samuel J. Huang<sup>1</sup>, Amanda Yung<sup>1</sup>,  
Jian Ding<sup>4</sup>, Oh-Sang Kwon<sup>1,3</sup>, Daphne Bavelier<sup>1,2</sup> & Dennis M. Levi<sup>4\*</sup>

<sup>1</sup>Department of Brain & Cognitive Sciences, University of Rochester, Rochester, NY 14627-0268, USA

<sup>2</sup>Faculty of Psychology and Education Sciences, University of Geneva, Switzerland

<sup>3</sup>School of Design & Human Engineering, UNIST, South Korea

<sup>4</sup>School of Optometry and Helen Wills Neuroscience Institute, University of California, Berkeley, CA, 94720, USA

<sup>†</sup>Work was performed while at the University of Rochester

<sup>\*</sup>Corresponding author

**Keywords:** Stereopsis; Virtual Reality; Perceptual learning; Strabismus; Amblyopia

## Summary

Stereopsis is the rich impression of three-dimensionality, based on binocular disparity—the differences between the two retinal images of the same world. However, a substantial proportion of the population is stereo deficient, and relies mostly on monocular cues to judge the relative depth or distance of objects in the environment. Here we trained adults who were stereo blind or stereo deficient due to strabismus and/or amblyopia in a natural visuo-motor task – a “bug squashing” game - in a virtual reality (VR) environment. The subjects’ task was to squash a virtual dichoptic bug on a slanted surface, by hitting it with a physical cylinder they held in their hand. The perceived surface slant was determined by monocular texture and stereoscopic cues, with these cues being either consistent or in conflict, allowing us to track the relative weighting of monocular versus stereoscopic cues as training in the task progressed. Following training most participants showed greater reliance on stereoscopic cues, reduced suppression and improved stereo-acuity. Importantly, the training-induced changes in relative stereo weights were significant predictors of the improvements in stereoacuity. We conclude that some adults deprived of normal binocular vision and insensitive to the disparity information can, with appropriate experience, recover access to more reliable stereoscopic information.

## 1. Introduction

Stereopsis is the impression of three-dimensionality—of objects “popping out in depth”—that most humans get when they view real-world objects with both eyes, based on binocular disparity, the differences between the two retinal images of the same world. However, a substantial proportion of the population is stereoblind or stereo deficient. The exact proportion depends on the specific test for stereopsis and the age of the subjects, but estimates of impaired stereopsis range from  $\approx 5$  percent [1]

\*Author for correspondence (dlevi@berkeley.edu).

1  
2  
3  
4 to as high as 34% in older subjects [2]. This impairment may have a substantial impact on  
5 visuomotor tasks, difficulties in playing sports in children and locomoting safely in older adults, and  
6 may also limit career options (see ref. 3 for a recent review).

7 Over the past five years there has been a renewed interest in restoring stereopsis in adults with  
8 strabismus, since the publication of “Fixing My Gaze”[4] in which Susan Barry, a neuroscientist,  
9 recounts her recovery from strabismus (a turned eye) and her amazement as she regains stereo-vision,  
10 and the description by Bruce Bridgeman, a vision scientist who had been stereo deficient all his life,  
11 of experiencing stereoscopic depth perception after viewing the 3D movie Hugo [5]. However, there  
12 are a limited number of experimental studies documenting recovery of stereopsis in adults who have  
13 long been deprived of normal binocular vision. Nakatsuka et al. [6] reported that adult monkeys  
14 reared with prisms had mild stereo deficiencies that improved through perceptual learning (PL) after  
15 10,000 - 20,000 trials. Astle, Webb & McGraw [7] reported on two cases of humans with  
16 anisometropic amblyopia whose stereopsis improved following a learning-based course of training,  
17 which included refractive adaptation followed by monocular PL as well as stereoscopic PL. Ding &  
18 Levi [8] provided the first evidence for the recovery of stereopsis through PL in human adults long  
19 deprived of normal binocular vision due to strabismus and/or amblyopia. They used a training  
20 paradigm that combined monocular cues that were perfectly correlated with the stereoscopic cues.  
21 Following PL (thousands of trials) with stereoscopic gratings, adults who were initially stereoblind or  
22 stereo-deficient showed substantial recovery of stereopsis. Importantly, these subjects reported that  
23 depth “popped out” in real life, and they were able to enjoy 3-D movies for the first time, similar to  
24 the experiences of Susan Barry and Bruce Bridgeman. Their recovered stereopsis is based on  
25 perceiving depth by detecting binocular disparity, but has reduced resolution and precision. Similar  
26 improvements were recently reported in a group of anisometropic and ametropic amblyopes who  
27 were trained with anaglyphic textures with different disparities [9].

28 How does training improve stereopsis? There are multiple cues to depth – both binocular (retinal  
29 disparity; convergence) and monocular (motion parallax, relative size, familiar size, cast shadows,  
30 occlusion, accommodation, texture gradient, linear perspective, aerial perspective, shading, lighting,  
31 and defocus blur). Stereo blind or deficient observers rely mainly on monocular cues. Ding & Levi  
32 [8] speculated that stereoblind or stereo deficient observers could learn to associate monocular and  
33 binocular cues to depth if they were highly correlated through repeated practice. In the current study  
34 we trained adult observers who were stereo deficient due to strabismus and/or amblyopia (lazy eye)  
35 in a natural visuo-motor task – a “bug squashing” game - in a virtual reality (VR) environment. The  
36 subjects’ task was to squash a virtual dichoptic bug on a slanted surface, by hitting it with a cylinder.  
37 The slant of the surface was determined by (i) purely stereoscopic cues (pure stereo-cue trials), or (ii)  
38 consistent monocular texture and stereoscopic cues (cue-consistent trials), or (iii) conflicting  
39 monocular texture and stereoscopic cues (cue-conflict trials). Importantly, our bug squashing training  
40 involves integrating not just multiple visual cues, but also the rich information from tactile and  
41 kinesthetic feedback. We hypothesized that training with multiple cues to depth with rich feedback  
42 might enable stereoblind or stereo deficient observers to increase their reliance on stereoscopic cues.  
43 Following training these observers showed increased reliance on stereoscopic cues (relative to  
44 monocular cues), reduced interocular suppression and significantly improved stereoacuity.

45 Our results have important implications for the recovery of visual function late in life, well  
46 outside the childhood period, until recently thought to offer the only real scope for plasticity. More  
47 broadly, they demonstrate the use of virtual reality as a promising approach for perceptual training of  
48 all kinds.  
49

## 50 2. Methods

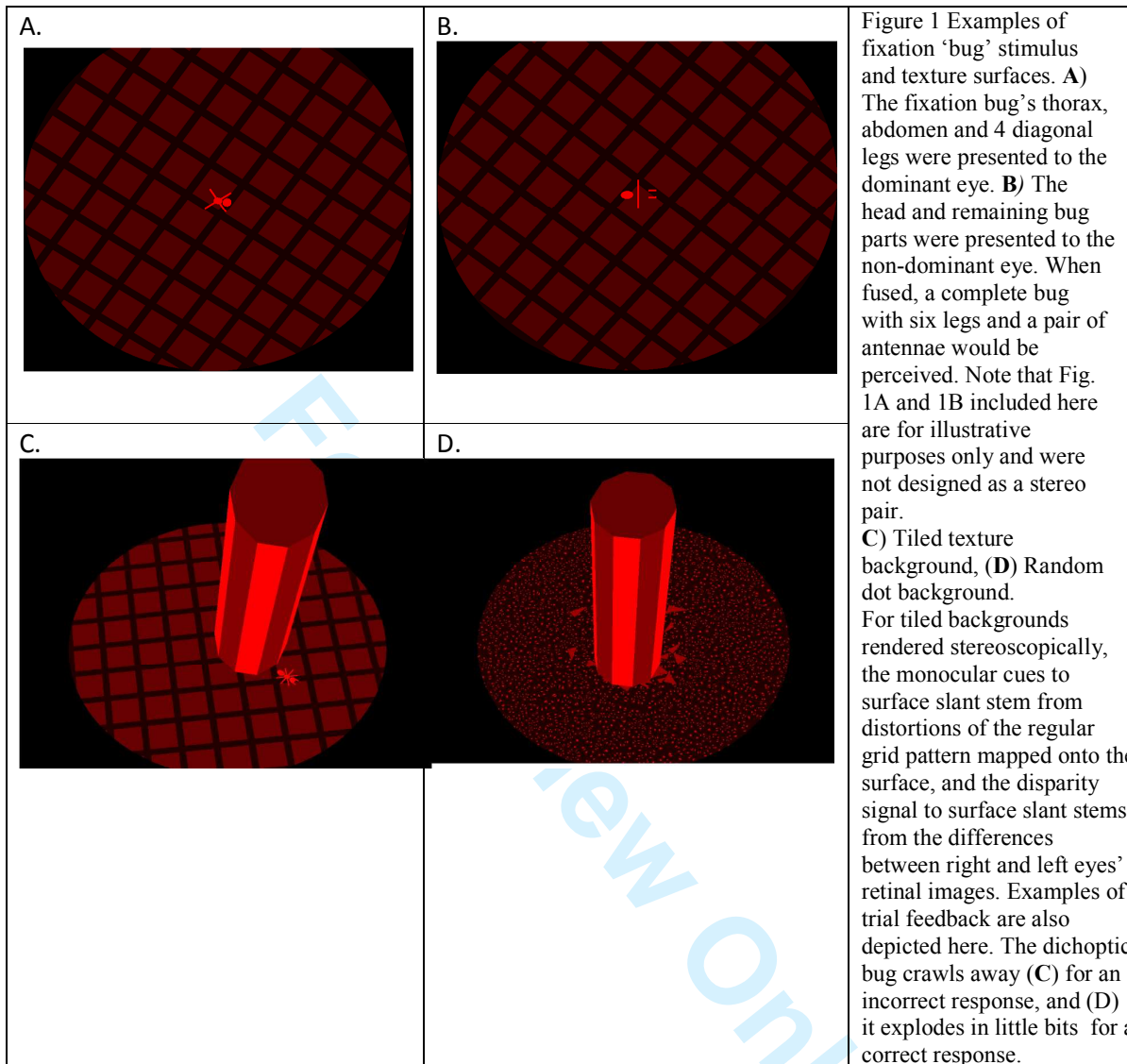
### 51 (a) Participants

1  
2  
3  
4 Eleven adults (mean age 34.7 years, range 19-56 years) with long-standing abnormal  
5 binocular vision completed the training study. Informed consent was obtained from all subjects  
6 conforming to the guidelines of the Research Subjects Review Board at the University of Rochester.  
7 Subjects were recruited mainly through referrals from local eye doctors and through print  
8 advertisements, and were paid \$10/hour for study participation. The inclusion criteria for the  
9 experimental group were (a) impaired stereopsis associated with one or more of the following  
10 conditions – anisometropic amblyopia, strabismic amblyopia, mixed (both anisometropic and  
11 strabismic) or pure strabismus, i.e., without amblyopia (b) normal ocular and general health, and (c)  
12 no history of eye surgeries except for those to correct strabismic deviation. Subjects with non-  
13 comitant and/or large angle strabismus (>30 prism diopters) were excluded. Anisometropia was  
14 defined as  $\geq 1D$  difference in spherical equivalent refraction between the two eyes. Those with  
15 manifest ocular deviation (strabismus), as indicated by the cover test, were classified as pure  
16 strabismics, and those having both anisometropia and strabismus were classified as mixed etiology.  
17 The clinical details of subjects are summarized in Table 1. Nine adults with normal acuity and  
18 binocular vision were recruited to provide normal control stereo weight data at baseline before any  
19 training. Three of these participants were then entered in the training phase with two completing 30  
20 training sessions and one completing 20 training sessions. Mean results for these 3 normal control  
21 observers are shown in figures 3, 5, 6, 7 and 8.  
22

23  
24 Subjects completed the training study in five phases - screening, pretesting, training, post-  
25 testing, and follow up. In the screening phase, subjects received a complete eye exam to determine  
26 whether or not they met the study inclusion criteria. Qualified subjects returned to the lab to complete  
27 a pre-training test battery (pretesting) that included baseline assessment of stereopsis, suppression,  
28 vergence and visual acuity. They then underwent a virtual reality based stereo training program  
29 (described below) for 35 sessions distributed over 8-11 weeks. The pre-training test battery was re-  
30 administered a few days after training to assess any training related changes (post-testing), and for  
31 the third time after a 2-month period of no intervention to assess retention effects (follow- up).  
32

### 33 34 (b) Training

35  
36 The visual stimuli consisted of textured, slanted virtual disks with a central fixation target  
37 presented in a virtual reality display (described in detail in ref. 10). Subjects wore Crystal Eyes  
38 shutter goggles during the training to view the stimuli (StereoGraphics Corporation, San Rafael, CA).  
39 All stimuli were drawn in red to minimize inter-ocular crosstalk by utilizing the relatively fast red  
40 phosphor of the monitor. A small dichoptic bug was rendered in the plane of the disk and served as  
41 the fixation target (Fig. 1 A & B). In the absence of suppression, and with accurate vergence, subjects  
42 perceive a complete bug with six legs and a pair of antennae. If the non-dominant eye is suppressed,  
43 the observer sees only the bug's thorax, abdomen and 4 diagonal legs. If the dominant eye is  
44 suppressed, the observer sees only the head and remaining legs. We used two types of textures to  
45 render the disks, so as to create stimuli with and without effective monocular cues to slant in depth -  
46 regular tiled textures, for which subjects with normal binocularity assign nearly equal weights to  
47 monocular and stereo cues (Fig. 1C), and randomly shaped dot textures (Fig. 1D) for which the  
48 texture cues are relatively uninformative [10, 11]  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



The subjects’ task was to squash the virtual bug by hitting it with a plexiglass cylinder measuring 6.4 cm in diameter and 12.7 cm in height and weighing 227 g. Subjects grabbed the cylinder and moved it from a starting plate positioned to the right of the virtual image (see Fig. 2), and were required to place it as flush as possible with the target surface over the bug. A three-camera Optotrak 3020 motion-capture system (Northern Digital Inc., Waterloo, Canada) was used to track the four infrared markers on the cylinder so as to estimate its 3D position and orientation in real-time. A virtual cylinder was rendered co-aligned with the real cylinder as it moved within the workspace. The cylinder was rendered with perspective and stereoscopic cues. As described below, since subjects viewed the stimuli through circular apertures to eliminate contextual cues provided by the monitor, the virtual cylinder was visible to the subject only when it approached the bug (typically for about ~250-300 ms before impact). When the cylinder was not within the subject’s field of view, the only information available about the orientation of the cylinder was proprioceptive information. Every trial was programmed to end when the cylinder contacted the target plate, which provided subjects with haptic feedback about the target slant.



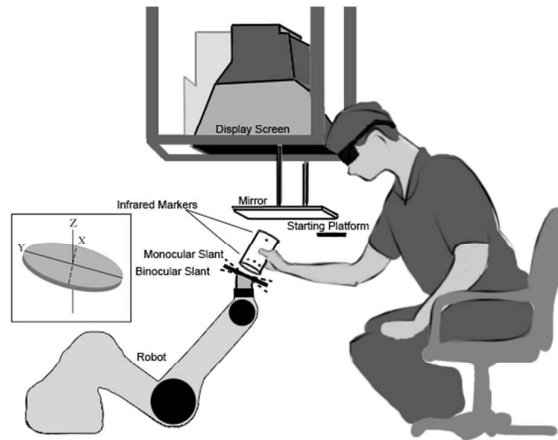


Figure 2: Stimuli were displayed on an inverted 22-inch Mitsubishi Diamond Pro 2070SB CRT monitor, with screen resolution 1152 x 864 pixels and refresh rate 120 Hz. Observers viewed the reflections of the stereoscopically presented stimuli through a mirror using Crystal Eyes shutter glasses (StereoGraphics Corporation, San Rafael, CA). An opaque black plate was placed beneath the mirror so that subjects could see only the virtual image of the display formed below the mirror. A PUMA 260 robot arm, invisible to the subject, coaligned a circular metal (target) surface with the virtual image at a distance of 63.5cm from the viewer's eyes. The slant of the target surface was defined as its orientation around the x-axis relative to subject's line of sight (see dashed line in the figure inset), and a zero degree slant indicated frontoparallel. Monocular and binocular slants differed on cue-conflict trials, but subjects perceived the stimuli as a single slanted disk. Head movements were restricted with head and chin rests. Subjects' task was to move a cylinder from the starting platform and place it flush onto the real surface to squash the virtual image. The slant of the real surface was determined by the mean of the monocular and binocular slants. The starting platform was 40 cm to the right of the target surface, 20 cm closer to the subject than the target surface, and 16.5 cm above the target surface (all measured from the subject's point of view).

At the beginning of each session, subjects performed a calibration procedure to estimate the positions of their eyes relative to the monitor to ensure accurate rendering of the virtual 3D space (see ref. 10 for details). The method of adjustment was then used to equalize the perceived contrast of the bug's half-images presented to the dominant and non-dominant eyes. Measurements were repeated six times, and the mean contrast was used to display the dominant eye's nonius bug during the training task. Prisms were used to optically align the nonius bug parts for experimental subjects with strabismus. No subject reported diplopic background textures post-alignment. Contextual cues from monitor edges and other surrounding objects were removed by placing an adjustable circular aperture before each eye that restricted the viewing angle to 11.9°.

Participants completed 35 sessions over a period of 8-11 weeks with an average of 3 sessions per week in the training phase. The first three usable sessions (pre) and last two sessions (post) were used to estimate changes from pre- to post-training stereo-weights. Two more sessions were run two months after training to estimate changes at follow up. In each session, target stimuli were rendered at five possible slants starting at 20 degrees and ending at 50 degrees away from fronto-parallel in 7.5-degree steps, with one of the 5 slants chosen randomly between trials. Sessions were run in blocks of 60 trials, with 6 blocks per session totaling 360 trials. 240 of these trials incorporated stimulus disks with tiled texture. In half of those trials, the monocular (tiled texture) and stereoscopic (disparity) cues to the surface slant were consistent (cue-consistent trials). In the other half, the slant specified by stereoscopic and texture cues differed by 7.5° (for example, monocular/stereoscopic slants: 42.5 / 35 or 27.5 / 35 – cue-conflict trials). The remaining 120 trials contained random dot textures that provided only stereoscopic cues (pure stereo-cue trials).

Each virtual reality training session began only after subjects reported that the dichoptically rendered bug halves were equally visible and aligned, ensuring fusion. Thereafter, each trial commenced with participants placing the cylinder squarely on the starting plate. This signaled the robot arm to orient a real surface, co-aligned with the slant specified for that trial. For cue-consistent trials, the slant of the real surface matched the slant suggested by (consistent) monocular and stereoscopic cues. For cue-conflict trials, the robot arm oriented the real surface at the average of the slants specified by the monocular and stereoscopic cues. A slanted virtual disk with the bug was then displayed for 2 seconds. At the end of 2 seconds, the bug spun 360 degrees in 250 ms (one full turn) to indicate the 'go' signal. Subjects were given 1.3 seconds after the go signal to move the cylinder from the starting plate and place it flush on the target surface to squash the virtual bug. Subjects were instructed to squash the bug only if both halves of the dichoptic bug were visible and aligned. The feedback strategy was identical for both cue-conflict and cue-consistent trials. If the orientation of the cylinder was within five degrees of the orientation of the target surface, the bug exploded. If not, the bug quickly ran across the surface and disappeared. When subjects took longer than 1.3 seconds, the trial was discarded and a message was displayed instructing the participants to respond faster. The discarded trials were randomly administered later in the block. The same procedure was repeated until the entire block was completed. Likewise, if subjects moved the cylinder prior to the go signal, the computer aborted the trial, displayed an error message, and repeated the trial later in the same block. For both cue-consistent and cue-conflict trials, in order to avoid a participant being able to tell where the robot arm moved based on auditory cues, the robot arm was made to move to two different slant angles before settling on the slant specified for that trial. The two slants were randomly chosen from 10-60°.

After a few initial training sessions, subjects took roughly 7-8 minutes to complete a block. Including calibration, each session lasted 60-75 minutes. Video game-like scores were given for correct and incorrect responses to motivate subjects. In order to make sure that subjects were not suppressing during the task and were accurately withholding movements on trials in which they did not perceive a bug, an additional 16 trials per block were included as 'no go' trials, that contained only one half of the nonius bug (presented to the dominant eye only) for fixation. Performance on these trials was, however, not recorded and will thus not be discussed further.

Nine subjects with normal vision provided baseline measurements for this task by participating in the first 3 sessions of the VR task. Three of these subjects continued with the VR task, one of them for 20 sessions and the other two for 30 sessions. Although these subjects were just run to pilot the VR task, we report below their data for qualitative comparison with the stereo-deficient patients.

### (c) Analysis of VR Training Data

In order to measure the influence of stereoscopic cues on the surface slant estimate used by subjects to plan their hitting movements, we regressed the slant of the cylinder just prior to contact with the surface (three optotrak frames - 24 msec. - prior to contact) against the slant depicted by the texture on the surface and the slant depicted by stereoscopic disparities using the equation –

$$s_{cyl} = w_{mono} s_{mono} + w_{stereo} s_{stereo} + k,$$

where  $s_{cyl}$  is the slant of the cylinder just prior to making contact with the surface,  $s_{mono}$  is the slant suggested by texture cues and  $s_{stereo}$  is the slant suggested by stereo cues. We normalized the weights to obtain a measure of the relative weight that subjects give to stereo to plan their movements

$$Relative\_Stereo\_Weight = \frac{w_{stereo}}{w_{stereo} + w_{mono}}.$$



Only trials containing regular tiled texture, and therefore effective monocular cues, were used for the regression. Trials were further limited to those containing slants between 27.5 and 42.5 degrees to minimize any effects on the regression that non-linearities in the mapping between stimulus and perceived slant might have. The cue-consistent trials were pooled together with the cue-inconsistent trials to determine the regression coefficients.

#### (d) Behavioural assessments

Suppression: Subjects' adjusted the relative contrast of the dominant eye's stimulus (the half bug) to match the perceived contrast of the non-dominant eye's half bug before each training session. We use the interocular contrast ratio (ratio of non-dominant to dominant eye contrast) that appears equal as a measure of interocular suppression (a ratio of 1 = no suppression; high values = strong suppression).

Stereoacuity: We used a standard clinical stereo-test to evaluate changes in stereoacuity – the Randot Circles Stereotest (Stereo Optical Co., Inc. – see ref 12 for details). Since this test contains monocular cues, we also included the Pure Disparity Test (PDT using 1 cpd sine wave grating stimuli) described by Ding and Levi [8], which contains no monocular cues.

Vergence Control Test: Vergence instability may negatively impact subjects' stereo-performance. We developed a novel psychophysical test to track the effect of training on vergence accuracy. Test stimuli consisted of two thin vertical lines presented either monocularly or dichoptically using Crystal Eyes shutter goggles. These were surrounded by a fusion frame consisting of 4 small wedge shaped markers with a central fixation dot shown binocularly at 0 disparity. Stimuli were viewed from a test distance of 1.5 meters. Phase 1 of the test required subjects to adjust the contrast of the dominant eye's stimulus (method of adjustments) so that it perceptually matched that of the non-dominant eye. Phase 2 generated a baseline measure of monocular line-alignment acuity of the dominant eye by presenting the test lines monocularly (using the contrast estimated in phase 1). Subjects were instructed to determine if the target (top) vertical line was displaced to the right or left with respect to the bottom (reference) vertical line, and respond accordingly by right or left mouse clicks. A random horizontal jitter was applied equally to both vertical lines in every trial, to prevent the use of the central fixation point as a reference for target displacement. Four interleaved staircases were constructed using 3-right / 1-left, 1-right / 3-left, 5-right / 1-left and 1-right/5-left rules. Phase 3 of the test required subjects to perform the same task as described in Phase 2, but using dichoptically presented lines and a binocularly presented fusion frame. This measure of dichoptic nonius alignment is similar to that described by McKee & Levi [13]. The target line was presented to the dominant eye and the reference line was presented to the non-dominant eye. Additionally, as a suppression check, subjects were instructed to click the mouse wheel if they only saw one line. Three consecutive 'suppression' responses warranted readjusting the line contrasts to the dominant eye, and the experiment was reset. A total of 600 trials were run.

Cumulative Gaussian psychometric functions were fit to the monocular and dichoptic test data. The variance of vergence noise was estimated from the standard deviation parameters of the cumulative Gaussian psychometric function fits for the monocularly and dichoptically presented line stimuli,

$\sigma_{vergence} = \sqrt{\sigma_{dichoptic}^2 - \sigma_{monocular}^2}$ . The 50% point on the psychometric function reflects bias or the constant shifts in eye alignment.

#### (e) Statistical Analyses

Data analyses were performed using the SPSS package. To test the effect of time, we used one-way

1  
2  
3  
4 repeated measures ANOVA including the three time points, pre-test, post-test and follow up. These  
5 were followed by Bonferroni adjusted pairwise comparisons to better identify where the specific  
6 differences lie.

7 Some of the measurements violated the normality assumption however; this was the case for  
8 suppression, randot, and vergence measurements. For those, the Friedman statistical test was applied  
9 This test is the non-parametric alternative to the one-way repeated measures ANOVA. Wilcoxon  
10 Signed-ranks tests were conducted to identify where the specific differences lie, with Bonferroni  
11 correction to control for inflation of type I error. All p values reported are two-tailed, except in pair-  
12 wise tests as mentioned.  
13

### 14 15 3. Results

#### 16 17 (a) Training increases the accuracy of performance.

18  
19  
20 Training resulted in improved accuracy of slant judgments for both dot stimuli (slant specified by  
21 stereoscopic cues only) and for textured stimuli containing either consistent monocular and  
22 stereoscopic cues, or conflicting monocular and stereoscopic cues, for both normal and stereo-  
23 deficient observers. Fig. 3 shows the mean 'hit rate' over sessions (left column) for the 11 stereo-  
24 deficient (top) and for the 3 stereo-normal observers (bottom). A trial was considered a hit trial, if  
25 the orientation of the cylinder was within plus or minus five degrees of the orientation of the target  
26 plate. The mean 'hit rate' was the average hit rate across all 5 slants.  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

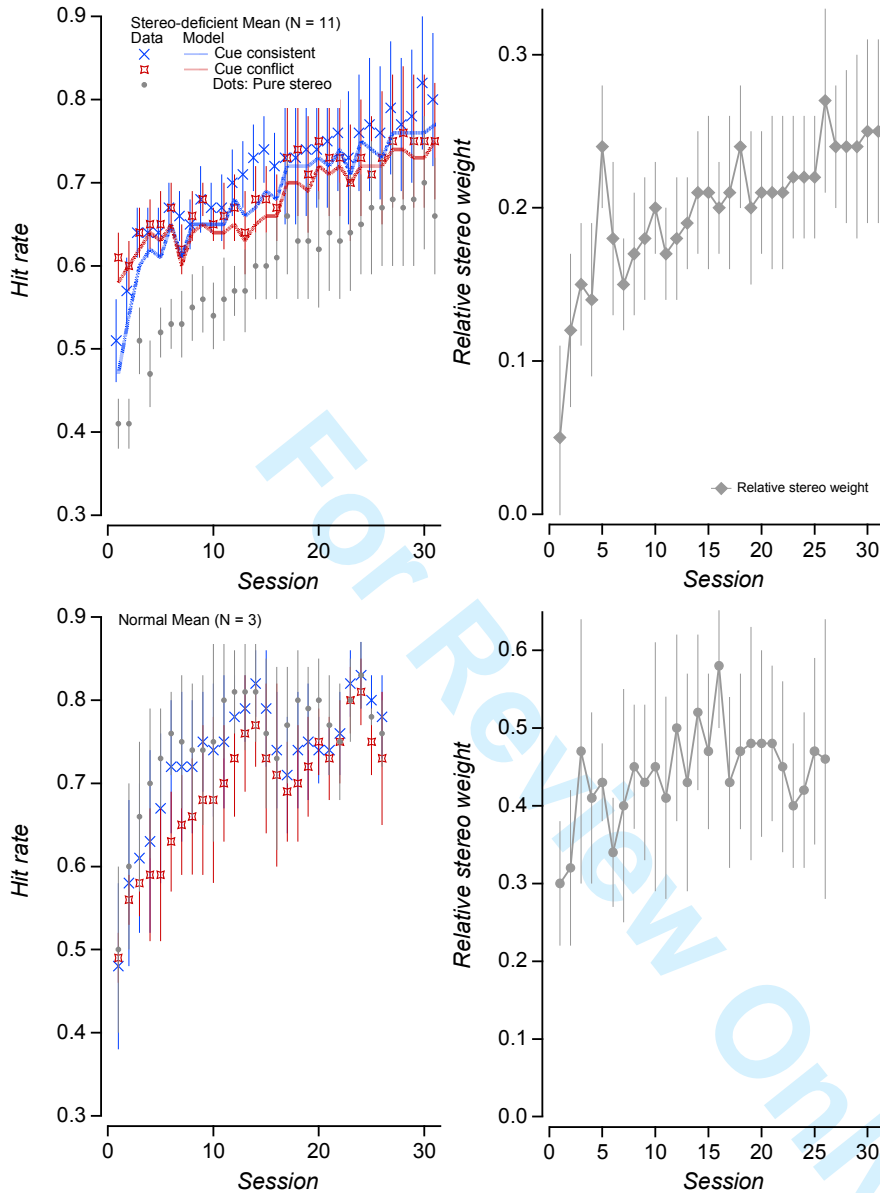


Fig. 3. Accuracy (left) and relative stereoweights (right) as training progresses. **Left column:** Mean ‘hit rate’ for the 11 stereo-deficient (top) and for 3 stereo-normal observers (bottom). A trial was considered a hit trial, if the orientation of the cylinder was within plus or minus five degrees of the orientation of the target surface (the plate), which was set to be at the mean orientation of the stereoscopic and monocular cues. The lines are the model output with the parameters shown in Fig. 4 (see the text). **Right column:** The mean of the relative stereoscopic weights is plotted for the 11 stereo-deficient (top) and for 3 stereo-normal observers (bottom; see section C of the data analysis for how these weights were derived).

In order to understand the improvement in performance, we fitted a simple Bayesian model to the mean accuracy data. The model has four parameters: stereoscopic cue noise, monocular cue noise, motor noise and a constant bias ( $k$ ). Sensory and motor noise are assumed to follow a zero mean Gaussian distribution with variances  $\sigma_{stereo}^2$ ,  $\sigma_{mono}^2$ , and  $\sigma_{motor}^2$ . We assume that subjects have a Gaussian prior distribution of surface slants whose variance reflects the actual variance of stereoscopic and monocular cues presented in the task ( $\sigma_{prior} = 9.2$  deg) [14], and integrate this prior knowledge with the stereoscopic and the monocular sensory signals in a statistically optimal fashion. Motor noise and bias are then added to the integrated sensory estimation. The resulting slant of the cylinder predicted by the model follows a Gaussian distribution whose mean ( $\mu_{\theta}$ ) and variance ( $\sigma_{\theta}^2$ )

can be computed as follows:

$\mu_{\theta} = w_{sense} (w_{stereo} \theta_{stereo} + (1 - w_{stereo}) \theta_{mono}) + k$ , where

$$w_{sense} = (\sigma_{prior}^2 \sigma_{stereo}^2 + \sigma_{prior}^2 \sigma_{mono}^2) / (\sigma_{mono}^2 \sigma_{stereo}^2 + \sigma_{prior}^2 \sigma_{stereo}^2 + \sigma_{prior}^2 \sigma_{mono}^2),$$

$$w_{stereo} = \sigma_{mono}^2 / (\sigma_{mono}^2 + \sigma_{stereo}^2),$$

$\theta_{stereo}$  is the slant of the surface in stereoscopic cue, and

$\theta_{mono}$  is the slant of the surface in monocular cue.

$$\sigma_{\theta}^2 = w_{sense}^2 (\sigma_{mono}^2 \sigma_{stereo}^2) / (\sigma_{mono}^2 + \sigma_{stereo}^2) + \sigma_{motor}^2$$

As can be seen in Figure 4, the stereo-deficient observers show a substantial reduction in stereoscopic cue noise over the course of training, a much smaller reduction in monocular cues noise, and essentially no change in the motor noise or bias parameters. The blue and red lines in Fig. 3 (top left column) show the simulated change in hit rate based on these model parameters.

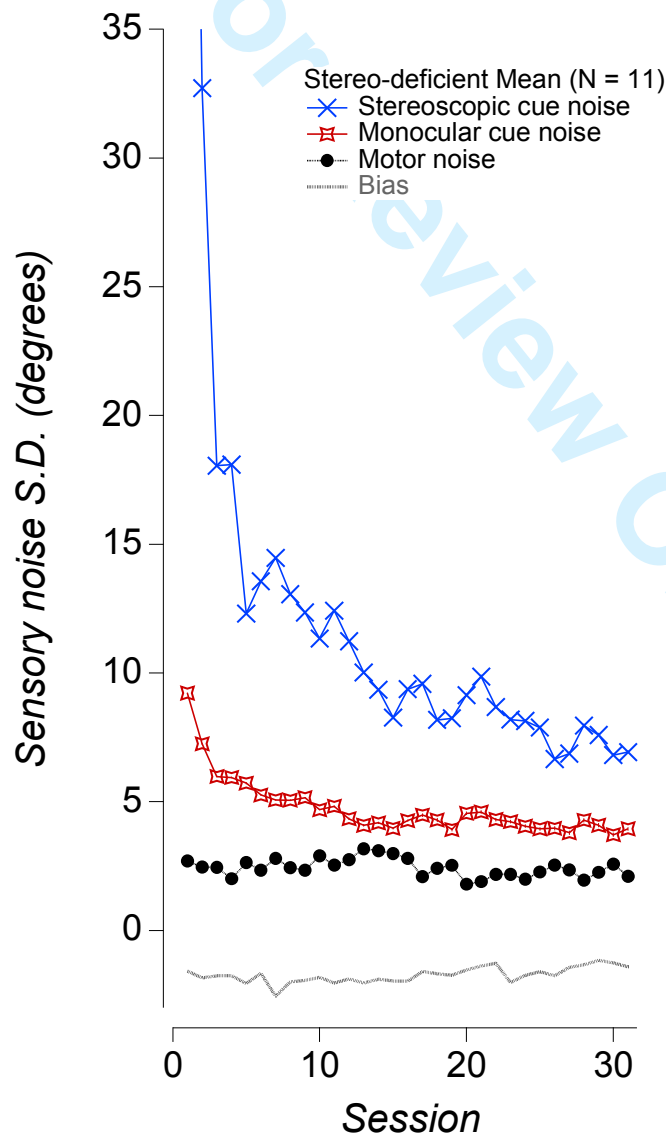


Fig. 4 Output of a simple Bayesian model fit to the mean accuracy data in the 11 stereo-deficient subjects. As participants trained in the task a significant reduction in stereoscopic and monocular cue noise is observed with motor noise and bias remaining stable throughout training. Thus, participants' estimate of depth cues became more reliable with training.

(b) Training increases the weighting of stereo cues.

The principal outcome measure, derived from the bug squashing data, is the relative stereo weight used by subjects to plan their hitting movements for stimuli containing regular textures. For these stimuli, the nine normally sighted control observers had strong relative stereo weights ( $0.39 \pm 0.05$ ) when tested under binocular conditions. Four of these subjects were retested with the dominant eye patched. Under these monocular conditions, their stereo weights were negligible ( $0.03 \pm 0.02$ ).

When first put on the task, our stereo deficient subjects were quite poor at the task as measured by the error in their hitting movements and a strong regression to the mean slant. One to six practice sessions were therefore used, depending on when accuracy and bias reached an asymptote. The number of practice sessions varied per subject, but data from the first three sessions with relatively stable performance were used to compute the pre-training weights. Owing to reduced variability in weight estimates with training, post-training (and follow-up) weights were based on data derived from two sessions after training.

The relative stereo weights of our 11 stereo-deficient subjects increased by, on average, a factor of  $3.07 \pm 0.85$  (Fig. 3 – top right panel shows the mean data). In contrast, the three stereo-normal control observers showed only a factor of  $\approx 1.3$  change in relative stereo weights as training proceeded (see Figure 3, bottom right panel) – from  $0.35 \pm 0.05$  (average of the first 3 sessions) to  $0.46 \pm 0.01$  (average of the last 2 sessions). It is interesting to note that the feedback given in the experiment would have provided a signal not to weight the stereo cue any higher than 0.5.

A one-way repeated measures omnibus ANOVA on log-transformed relative stereo weights (to offset violations of normality) was conducted to compare group performance at pre, post, and follow up. The results showed that subjects assigned significantly more weight to stereoscopic cues after training [ $F(2, 20) = 7.77, p = 0.003$ ]. Bonferroni adjusted pairwise comparisons showed increased stereo weights (decreased monocular cue weights) after training ( $p = 0.004$  one tailed, mean change =  $0.15 \pm 0.04$ ), with the difference being still seen at follow up ( $p = 0.028$  one tailed, mean change =  $0.13 \pm 0.04$ ), indicating that the improvements were largely retained after two months of no intervention.

Fig. 5 shows the post vs. pre-training (solid symbols) relative stereo weights for each of the 11 stereo-deficient observers. Points above the unity line indicate an increase in relative stereo weights following training. Eight of the eleven subjects showed a numerical increase in stereo weights following training, with these improvements being largely maintained at follow up, 2 months later (except for two subjects who showed regression in gains). The inset in Fig. 5 shows the evolution of this increase in relative stereo weights for a strabismic subject (S1) who initially showed very low weighting of the stereo cues. A bootstrap analysis (10,000 iterations), performed to assess the intra-individual changes in relative stereo weights after training compared with their respective baselines indicates that the upweighting of the stereoscopic cues following training was statistically significant for six of the eleven stereo-deficient subjects at post-training (all  $p$ s  $< .001$ ). In addition, two of them (M2 and S4) showed the right trend from pre to post-training to follow up, but failed to reach significance in part due to the high variance of the pre-test estimates.

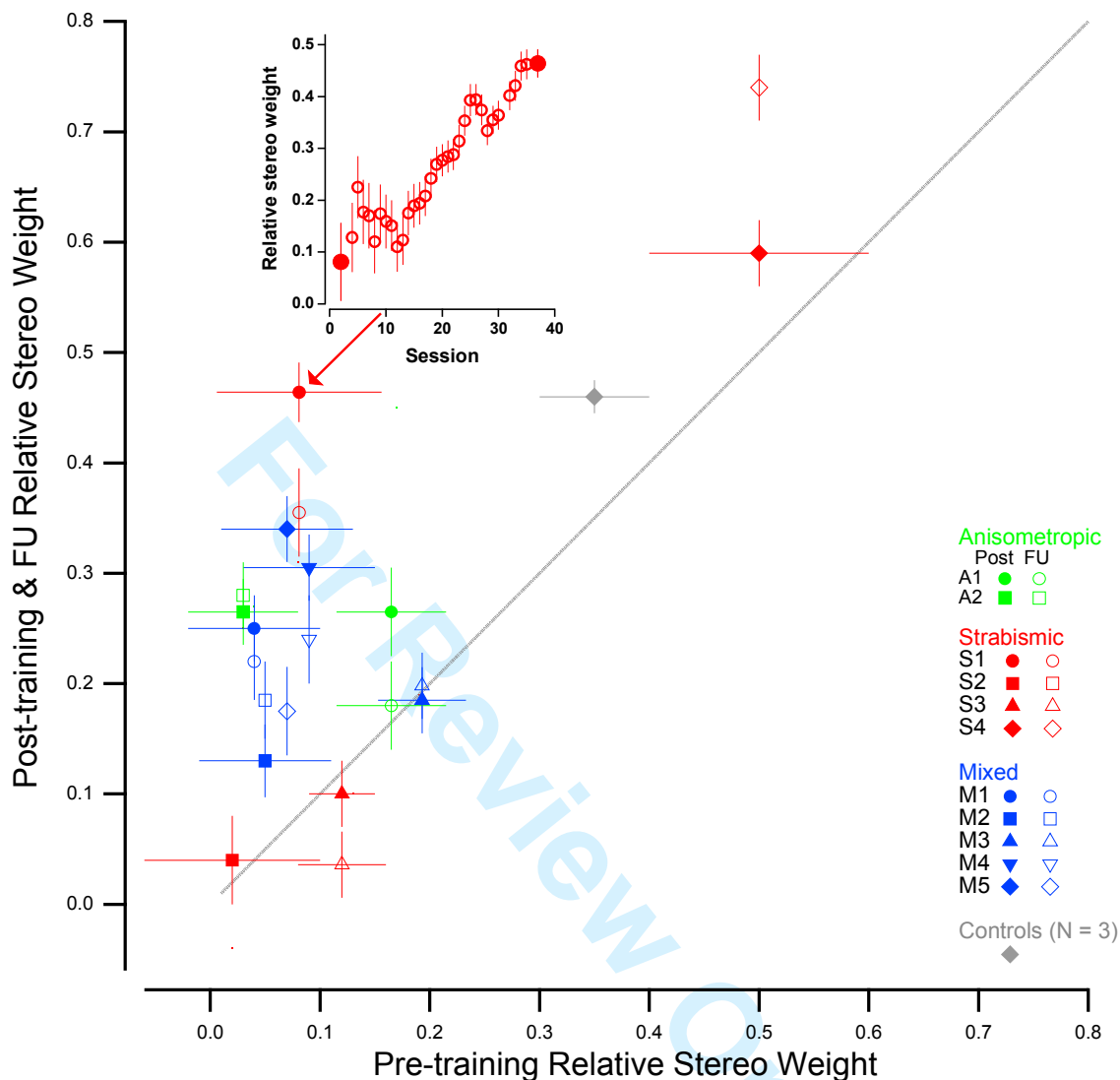


Figure 5. Post vs. Pre-training (solid symbols) and Follow Up (open symbols) relative stereo weights. Each colored symbol shows the data of a single stereo-deficient observer. The gray diamond near the abscissa shows the pre- and post-training mean relative stereo weights of the 3 normal control observers who underwent training. Data above the diagonal unity line indicate increases in relative stereo weights. The inset shows the evolution of the increased relative stereo weights for strabismic subject S1 (indicated in the main figure by an arrow). Solid symbols in the inset represent the pre-to-post training weights and the open symbols the pre-to-follow up training weights.

## (b) Behavioural measures of suppression, stereoacuity and visual acuity following training.

### Suppression:

Suppression was substantially reduced following training. The Friedman statistical test was applied since the data were non-normally distributed. We found that the bug squashing training resulted in a significant change to levels of suppression (Fig. 6, Friedman test,  $\chi^2(2, N = 11) = 13.35$ ,  $p = 0.001$ ). Specifically, reduced suppression was found in all but one subject after training (Wilcoxon signed-ranks test,  $Z = -2.67$ ,  $p = 0.0025$  one tailed,  $r = -0.57$ ; Median change in interocular contrast ratios = 0.44). These effects were retained at follow up compared to baseline ( $Z = -2.67$ ,  $p = 0.0025$  one tailed,  $r = -0.57$ ; Mdn change = 0.35), with no significant difference between



the magnitudes of suppression measured at post-training vs. at follow up ( $Z = -1.46$ ,  $p = 0.12$  one tailed,  $r = -0.31$ ; Median change = 0.0).

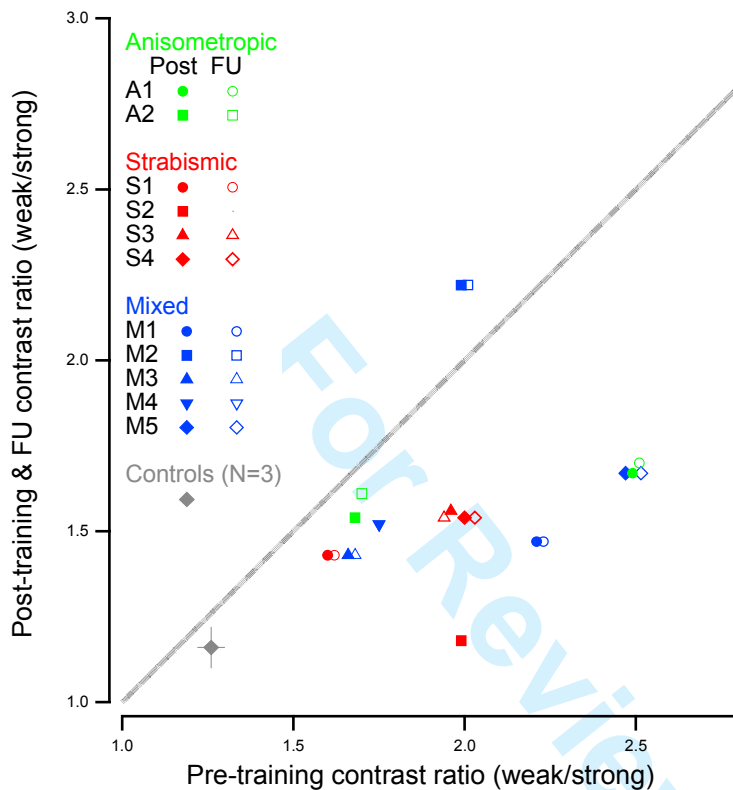


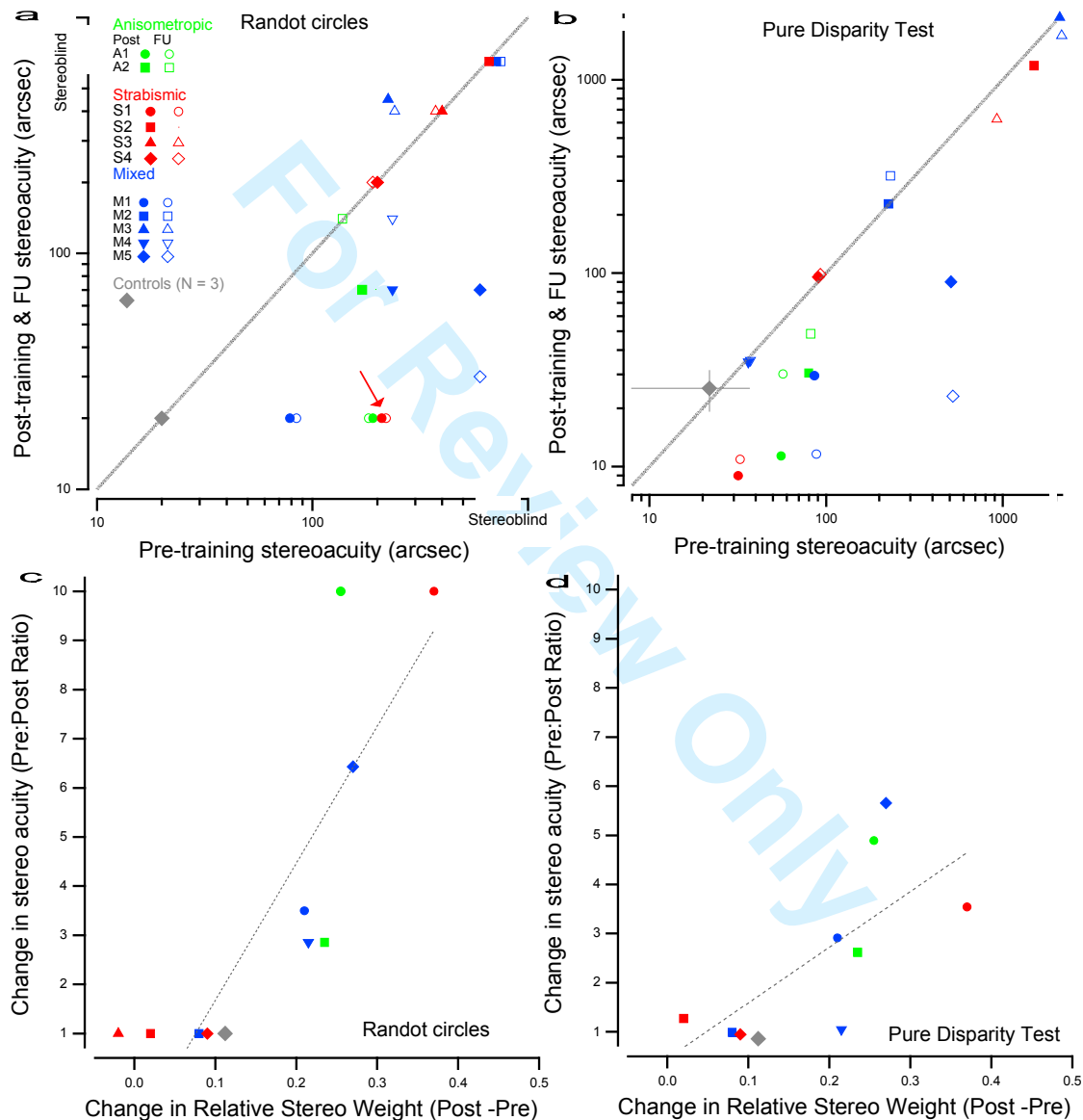
Figure 6. The effect of training on suppression. Pre vs. Post-training (solid symbols) and Follow Up (open symbols) interocular contrast ratio (weak/strong). Each colored symbol shows the data of a single stereo-deficient observer. The gray diamond shows the pre- and post-training mean data of the 3 normal control observers who underwent training. Data below the diagonal unity line indicate reduced suppression. A horizontal offset has been applied to avoid symbols (post and FU) from overlapping.

#### Stereoacuity:

Importantly, bug squashing led to improvements not only on the trained task, but also in the untrained stereo acuity tasks. All six subjects who showed significant increases in relative stereo weights at post-test, also demonstrated an improvement in stereoacuity as measured by the Randot Circles Test, (Fig. 7a solid symbols), and five of these six subjects also showed significant improvements in stereoacuity (all  $p$  values  $< 0.01$ ; 1000 bootstrap iterations were performed to assess intra individual differences in pre- and post stereothresholds) on the Pure Disparity Test (Fig. 7b). These improvements were largely maintained at follow up two months after the cessation of training (Fig. 7 b - open symbols), with the exception of two anisometropic amblyopes who either discontinued use of contact lenses post-training (A2), or used them only sparingly (A1). These two subjects demonstrated some regression of stereoscopic gains at follow up. Finally, the two subjects who showed the right numerical trend in relative stereo weights as the experiment proceeded (M2 and S4) did not show stereoacuity transfer as measured either with the Randot or with the Pure Disparity test.

*Randot Circles Test:* Statistical tests were run on log stereo-thresholds. A value of 2.78 log arc seconds (or 600 arc seconds) was arbitrarily assigned if subjects failed the Randot Circles test. Subjects unable to perform at that largest disparity were labeled as 'stereoblind'. There was a significant change in stereoacuity over time, Friedman test,  $\chi^2(2, N = 11) = 6.64$ ,  $p = 0.031$ . Post hoc comparisons showed that subjects exhibited a trend for improved stereoacuity after training (Wilcoxon,  $Z = -1.866$ ,  $p = 0.039$  one-tailed,  $r = -0.4$ ; Median change = 0.46 log arc sec, equivalent to 65% median improvement), and at follow up compared to pre-training baselines ( $Z = -2.106$ ,  $p = .02$  one-tailed,  $r = -0.45$ ; Median change = 0.15 log arc sec, equivalent to 30% improvement).

**Pure Disparity Test:** A one-way repeated measures ANOVA showed that stereoacuity as measured by the Pure Disparity Test improved significantly after training (main effect of time:  $F(2, 20) = 4.86$ ,  $p = 0.019$ ). Post hoc comparisons using Bonferroni correction showed that the stereoacuity improved post-training compared to pre-training baselines ( $p = 0.02$  one tailed; mean change =  $0.28 \pm 0.09$  log arc sec, equivalent to  $35.7 \pm 10.9\%$  improvement), and this improvement was largely retained at follow up ( $p = 0.046$  one tailed; mean change at follow up from pre-training =  $0.41 \pm 0.16$  log arc sec).



**Figure 7. Panel A- Randot Circles test and Panel B - Pure Disparity Test.** Pre vs. Post-training (solid symbols) and Follow Up (open symbols) stereo thresholds (Randot Circles test – Panel A; Pure Disparity Test – Panel B). Each colored symbol shows the data of a single stereo-deficient observer. The gray diamond is the mean pre/post data of the 3 control subjects who underwent training. Data below the diagonal unity line indicate decreases in stereo thresholds (i.e., improved stereoacuity). The red arrow in A indicates the subject (S1), with the largest change in stereo weights. Note that small horizontal shifts have been applied to some of the data points to avoid symbols overlapping. **Panel C- Randot Circles test and Panel D - Pure Disparity Test.** The change in stereo thresholds (Pre:Post ratio) vs the change in relative stereo weights (Post-Pre) for the Randot circles test (Panel C), and for the Pure Disparity Test (PDT) (Panel D), which has no monocular cues. The lines show the best fitting linear regressions for the Randot circles (solid line) and PDT (dashed line) respectively.

While these results show that learning a visuomotor task that focused on surface slant estimation transfers to stereo depth judgments, a stronger test for generalization would be the ability to predict improvements on the transfer tests from the post-training changes in stereo weights. Fig. 7 C and D plot changes in stereo threshold (Pre:Post ratio) as measured by the Randot Circles and Pure Disparity Test, respectively, against changes in stereo weights (post – pre relative stereo weights). The regression lines fit the data well, indicating that training-induced changes in stereo weights significantly predicted improvements on the Randot Circles ( $\beta = 0.75$ ,  $t = 3.44$ ,  $p = 0.007$ ;  $R^2 = 0.57$ ), and Pure Disparity Test ( $\beta = 0.9$ ,  $t = 6.0$ ,  $p < 0.0001$ ;  $R^2 = 0.8$ ).

#### Visual acuity:

We examined whether direct stereo training might also offer training-induced benefits to visual acuity (Fig. 8). Six of the eleven subjects showed an improvement in visual acuity immediately post training ( $> 0.04$  logMAR), with two of these subjects regressing and one lost at follow up. Of the three subjects that showed substantial ( $\geq 0.10$  logMAR) improvements, two showed no post-training improvements in stereopsis, suggesting changes in the non-dominant eye acuity are not tightly coupled to the training-related increases in relative stereo weights, in agreement with the results of Ding & Levi (2011). For the group-as-a-whole, there was a trend for improved visual acuity in the non-dominant eye after training compared to the pre-training baseline; however, this effect was only marginally significant (repeated measures ANOVA,  $F(2, 20) = 3.23$ ,  $p = 0.061$ ). Acuity of the dominant eye was unaltered post training ( $F(2, 20) = 0.72$ ,  $p = 0.5$ ).

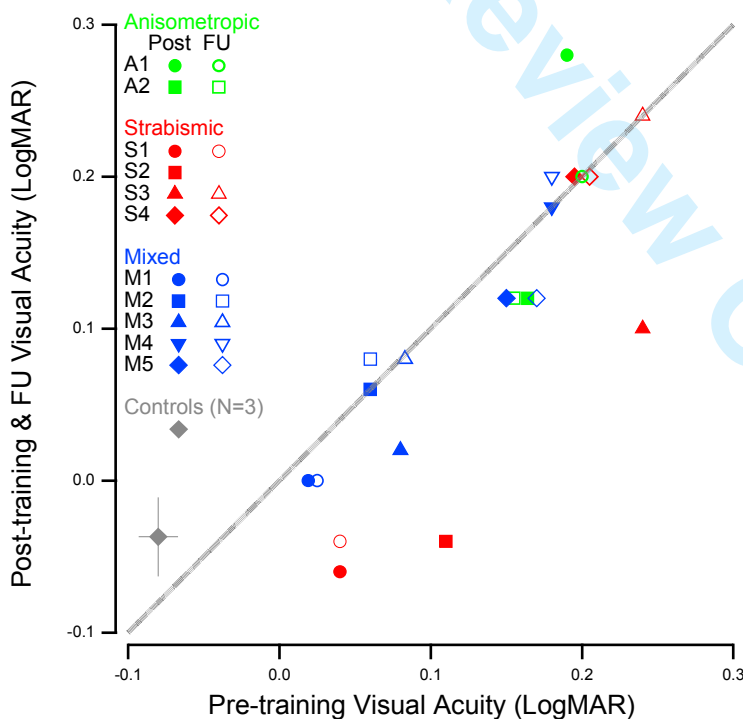


Figure 8. Pre vs. Post-training (solid symbols) and Follow Up (open symbols) visual acuity (LogMAR). Each colored symbol shows the data of a single stereo-deficient observer. The gray diamond shows the mean pre/post data of the 3 normal control observers who underwent training. Data below the diagonal unity line indicate increases in visual acuity.

#### Vergence:

We measured vergence noise in eight of the participants (Supplemental Figure S.1). Prior to bug squashing training, the vergence noise in the experimental group was 4.9 times higher ( $9.72 \pm 3.5$  arc minutes,  $n = 8$ ) compared to the normally sighted subjects ( $1.97 \pm 0.2$  arc minutes,  $n = 4$ ). For the stereo-deficient group as a whole, there were no significant differences in vergence noise measured pre-training, post-training and at follow up (Friedman test,  $\chi^2(2, N = 8) = 2.25$ ,  $p = 0.36$ ). Likewise, there were no changes in vergence bias with training (Friedman test,  $\chi^2(2, N = 8) = 3.25$ ,  $p$

1  
2  
3  
4 = 0.24). Four of these eight stereo-deficient subjects showed improvements in relative stereo weights  
5 following training, two of whom showed a post-training reduction of vergence noise, one showed an  
6 increase, and one showed no change. There was a lack of consistency in the direction of change post  
7 training. Thus there is insufficient evidence to indicate that changes in the vergence noise can explain  
8 the training-related improvements in stereo weights in these subjects. We acknowledge that our  
9 vergence measure is a subjective measure, and thus may not be sufficiently sensitive to detect very  
10 small changes in oculomotor control ( $\approx 1.5 - 2$  arc minutes in our normal control subjects –  
11 Supplemental Figure S1).  
12

#### 13 Determinants of stereo-weights improvements:

14  
15  
16 We also examined other potential contributing factors to post-training changes in subjects'  
17 relative stereo weights. First, changes in interocular suppression failed to predict changes in stereo  
18 weights ( $\beta = -0.2$ ,  $t = -0.65$ ,  $p = 0.5$ ;  $R^2 = 0.04$ ). Second, we investigated whether any of the pre-  
19 training measures – logMAR acuity (non-dominant eye), interocular difference in acuity, stereo  
20 weights, age, and Randot and PDT stereosensitivities – could predict training related changes in  
21 stereo weights. Multiple regression analysis revealed that only the initial Pure Disparity Test  
22 significantly predicted training-induced changes in stereo weights ( $\beta = -0.8$ ,  $t = -3.9$ ,  $p = 0.003$ ,  $R^2 =$   
23  $0.63$ ). None of the other measures did ( $p > 0.05$ ). Subjects with the best stereo acuity on the pure  
24 disparity tests at the outset, showed the greatest increases in relative stereo weights, while those with  
25 no measurable stereopsis showed no change.  
26  
27  
28  
29

## 30 4. Discussion

31  
32 While there have been a large number of studies of the effects of perceptual learning and  
33 videogame play in adults with amblyopia [for recent reviews see refs 15 -18], only a few studies have  
34 trained stereopsis directly (for a review ref 3). Our results show that training in a natural visuo-motor  
35 task in which some stimuli contained monocular texture as well as stereoscopic cues (cue-consistent  
36 trials), some contained only stereoscopic cues and some contained conflicting monocular and stereo  
37 cues, enabled adults with long standing deficits in stereo vision to upweigh their reliance on  
38 stereoscopic cues (relative to monocular cues). Following training participants showed improved  
39 accuracy for detecting slants in depth. Our simple Bayesian modeling indicates a very substantial  
40 reduction in stereoscopic cue noise over the course of training in stereo-deficient observers, along a  
41 more modest reduction in monocular cues noise. The dramatic increase in stereoscopic cue reliability  
42 as training progressed is consistent with the increased reliance of the observers on stereoscopic cues  
43 (i.e., increased relative stereo weights). In addition, stereo-deficient observers also showed reduced  
44 suppression, significant improvement in stereoacuity and a weak trend for improved visual acuity.  
45  
46  
47

48 Out of the 11 participants trained, 6 showed significantly greater reliance on stereoscopic cues  
49 as their training on our virtual reality bug-squashing game progressed, with an additional two  
50 showing the right numerical trend. Of the six who showed a significant effect, all showed improved  
51 stereo-acuity as measured by the randot at pre and post-training. Although only marginally  
52 significant, we note, as others have done before, that the Randot may lack the needed sensitivity to  
53 assess pre-to-post-test improvements. For example, strabismic subject S1 (shown by the arrow in  
54 Figs 5 and 7), achieved a stereoacuity of 20 arc sec., the lower limit of the Randot Circles test – this  
55 is the subject shown in the inset of Figure 5 whose relative stereo weights improved substantially  
56 over the course of training. We suspect that this floor (lowest tested value on the test) underestimates  
57 S1's stereoacuity, since he demonstrated a stereoacuity of  $\approx 9$  arc sec on the Pure Disparity Test.  
58 Five participants showed highly significant post training improvements on the Pure Disparity Test  
59  
60

(Table S1), and crucially, the best determinant of whether stereo-weights improved was performance on the Pure Disparity test at pre-test.

Use of different sensory cues:

Adults with normal visual experience reduce sensory uncertainty by integrating information from different modalities (e.g. touch and vision [19, 20]) and within the same modality (e.g. binocular disparity and texture information) for judging slant in depth [10, 21, 22]. However, observers deprived of normal binocular visual experience early in life have poor or absent stereopsis, and therefore rely on texture information for making judgments of surface slant. Thus, in our sample of observers with abnormal binocular vision, the pre-training relative stereo weight was, on average  $\approx 0.11 \pm .03$ . In contrast, normal control subjects had a relative stereo weight of  $0.39 \pm 0.05$ . Over the course of training, the relative stereo weights of the stereo-deficient subjects increased substantially. It is interesting to note that while the adult visual system is optimized for reducing sensory uncertainty by integrating texture and stereoscopic cues to slant, mature sensory integration is not evident until age 12 [23]. For infants and young children the limitation is not due to insensitivity to one cue but to an immaturity in sensory integration. In contrast, adults deprived of normal binocular vision are initially insensitive to the disparity information. With experience in bug squashing, the disparity information becomes more reliable. Importantly, our bug squashing training involves integrating not just multiple visual cues, but also the rich information from tactile and kinesthetic feedback [22].

Real life impact:

Following training two participants reported better distance judgement during driving, and one was able to appreciate depth from autostereograms for the first time. Finally, one important practical outcome of our study is the finding that strabismic subjects who demonstrated post-training stereo-improvements were able to perform the clinical Randot test without the use of prisms, despite the presence of uncorrected ocular deviation. Compensating the deviation with prisms did not further improve stereo acuity in any of our strabismic subjects. We hypothesized that if this finding were true, then the stereo weights should similarly be unaffected by lack of compensation for deviation. To assess this, we measured stereo weights without prisms for six subjects who were initially trained on the bug squashing task with prisms. Our results showed that the difference between relative stereo weights assessed with and without prisms was negligible (mean =  $0.01 \pm 0.02$ ).

Our strabismic subjects who recovered stereopsis had stereoacuties after training, as low as 9 arc sec (S1 on the PDT). How do we reconcile the improvement in stereopsis in the presence of an uncorrected strabismus? It has long been known that fusion is not a requirement for stereopsis. For example, observers with normal vision can correctly localize objects in depth even when they appear double [24]. Westheimer & Tanzman [25] demonstrated that observers with normal binocular vision can correctly identify the depth of diplopic stimuli with disparities up to about 7 degrees. Blakemore [26] reported that the largest disparity that supported signed depth judgments was  $\sim 12$  degrees. More recently, Dengler & Kommerell [27] showed that normal subjects could distinguish binocularly disparate images from monocular double images with the same angular separation over large distances (up to 21 degrees), where one eye's target was presented in the fovea, and the other eye's target was presented in the periphery. This is similar to the situation in strabismus, and, as Dengler & Kommerell speculate, these long-range connections between the fovea of one eye and the periphery of the other may be the basis for anomalous correspondence in patients with strabismus. One potential issue with these studies is that subjects can assign a 'correct' disparity to an isolated monocular target in one eye, as though it were matched to an invisible target in the fovea [28, 29]. Dengler & Kommerell, like the Foley et al. [30] study before it, did employ monocular targets as controls, where the subjects had to identify the monocular targets with a label "on". What is



1  
2  
3  
4 interesting is that many subjects did respond based on 'eye-of-origin' information, and so their data  
5 were discarded. Perhaps there were subtle differences that allowed some subjects to discriminate the  
6 targets that were supposed to be labeled 'crossed' from the targets that were supposed to be labeled  
7 'on'. However, the main point is that there is evidence for at least coarse stereopsis for unfused  
8 stimuli in normal vision. Coarse stereopsis is considered to be important to extend the range of  
9 disparity sensitivity, as a guide to vergence eye movements, and as a "back up" system for  
10 individuals with strabismus [31]. Indeed, there is recent evidence that coarse stereopsis develops  
11 much earlier than fine stereopsis [32], and that it may be relatively "spared" in individuals with a  
12 history of amblyopia [33]. Whether the same mechanisms are capable of supporting stereoacuity  
13 better than one arc minute in strabismic patients, or whether there are specialized mechanisms  
14 supporting anomalous correspondence is a matter for further research. For now, the present study  
15 documents that it is possible for strabismic patients to recover stereoacuity as low as 9 arc sec.  
16

17  
18 Why does recovery of stereopsis require heroic methods?  
19

20  
21 Isn't the rich real world stimulation enough? We believe that there are several reasons that  
22 the natural environment is not sufficient. First, under normal binocular viewing, stimuli in the two  
23 eyes do not have equal perceived contrast, leading to suppression of the weak eye by the strong eye  
24 [18, 34, 35]. Moreover, for strabismic subjects, the two eyes' images fall on non-corresponding  
25 retinal areas, precluding normal fusion and triggering suppression. Thus under normal viewing higher  
26 brain areas involved in sensory integration [36 - 39] receive weak and unreliable information from  
27 the weak eye, resulting in a down-weighting of disparity information. In contrast, in our bug  
28 squashing game, the stimuli to the two eyes were perceptually matched, by reducing the contrast  
29 presented to the strong eye and were carefully aligned, so that the images can be fused. By  
30 combining this new aligned and balanced visual input with a visuomotor task (bug squashing),  
31 providing monocular depth cues as a scaffold, and giving trial by trial force feedback, subjects can  
32 learn to attend to input from the amblyopic eye [40] and can learn the correlations between the  
33 "corrected" visual input and the depth of objects in the world. Finally, we should emphasize that our  
34 bug squashing training was conducted in a virtual reality environment enabling us to recreate "the  
35 natural way in which perception and action are intimately entwined with the environment" [41].  
36

37  
38 Our current results, taken together with previous studies (reviewed in ref 3), suggest that to  
39 optimize stereo recovery it is critical to initially provide aligned and balanced input to the two eyes.  
40 This may require substantial fusion training [8] prior to direct stereo training in a rich environment  
41 which combines a natural visuomotor task.  
42

## 43 5. Summary and Conclusions

44  
45  
46 We trained adults who were stereoblind or stereo-deficient on a natural visuo-motor "bug  
47 squashing" task in which some stimuli contained monocular texture cues as well as stereoscopic cues  
48 that were consistent with each other, some contained only stereoscopic cues and some contained  
49 conflicting monocular and stereoscopic cues. Following training, 8 out of 11 observers upweighted  
50 their reliance on stereoscopic cues (relative to monocular cues), while 6 of them showed reduced  
51 suppression and significant improvements in stereoacuity.  
52

53 Our results have important implications for the recovery of visual function late in life, well  
54 outside the "critical period", which until recently was thought to offer the only real scope for  
55 plasticity. More broadly, our approach demonstrates the potential power offered by virtual reality for  
56 perceptual training of all kinds.  
57  
58  
59  
60



## Additional Information

### Acknowledgments

Our dear friend and colleague David Knill died tragically last year. We, and indeed the entire vision community, will miss him sorely, and we dedicate this paper to his memory. We thank our research assistants and study coordinators Olga Pikul, Leslie Chylinsky, Aleksandra Fazlipour and Gary Volkell. We also thank Dr. Hu for assistance with programming. We thank Drs. Gearinger and DePaolis for referring patients to our study. Finally, this work could not have happened without the cooperation of our subjects.

### Ethics

The study was approved by the Institutional Review Board at the University of Rochester, and informed consent was obtained from all subjects.

### Data Accessibility

All manuscripts which report primary data (usually research articles) should include a Data Accessibility section which states where the article's supporting data can be accessed. This section should also include details, where possible, of where to access other relevant research materials such as statistical tools, protocols, software etc. If the data has been deposited in an external repository this section should list the database, accession number and link to the DOI for all data from the article that has been made publicly available, for instance:

DNA sequences: Genbank accessions F234391-F234402 (<http://dx.doi.org/xxxxx>)

Phylogenetic data, including alignments: TreeBASE accession number S9123 (<http://dx.doi.org/xxxxx>)

Climate data and MaxEnt input files: Dryad doi:10.5521/dryad.12311 (<http://dx.doi.org/xxxxx>)

If the data is included in the article's Supplementary Material this should be stated here, for instance:

The datasets supporting this article have been uploaded as part of the Supplementary Material.

### Authors' Contributions

The initial conceptual framework for the VR training was developed by DK. DK, DL and DB designed the study and obtained a research grant from the National Eye Institute. Adapting the VR training paradigm for clinical application: IV for design aspects and feedback, AY and DK for the programming. Piloting and fine-tuning of the training and vision assessments, primarily IV with contributions from AY, SH and DK; Running the Study: Primarily IV with contributions from SH and the supervision of DB; Data analysis: primarily IV with the help of DK and DB; modeling of setero-weight data: O-SK; JD developed the PDT and assisted with its data analysis. Writing: primarily IV and DL with the feedback of DK and DB, and other authors' contributing as needed. Final figures: DL.

DB and DL are co-senior authors.

### Competing Interests

We have no competing interests.

### Funding

This research was supported by grants RO1EY020976 from the National Eye Institute to D.L., D.K. and D.B.

## References

1. Tam, W. J., and L. B. Stelmach. (1998) "Display Duration and Stereoscopic Depth Discrimination," *Can J Exp Psychol*, vol. **52**, no. **1** 56-61.
2. Zaroff, C. M., M. Knutelska, and T. E. Frumkes. (2003) "Variation in Stereoacuity: Normative Description, Fixation Disparity, and the Roles of Aging and Gender," *Invest Ophthalmol Vis Sci*, vol. **44**, no. **2** 891-900.
3. Levi, D. M., D. C. Knill, and D. Bavelier. (2015) "Stereopsis and Amblyopia: A Mini Review," *Vision Res*, vol. **114** 17-30.
4. Barry, S.O. (2009). *Fixing My Gaze: A Scientist's Journey Into Seeing in Three Dimensions*. New York, NY: Basic Books.
5. Bridgeman, B. (2014) "Restoring Adult Stereopsis: A Vision Researcher's Personal Experience," *Optom Vis Sci*, vol. **91**, no. **6** e135-9.
6. Nakatsuka, C., B. Zhang, I. Watanabe, Zheng, H. Bi, L. Ganz, E. L. Smith, R. S. Harwerth, and Y. M. Chino. (2007) "Effects of Perceptual Learning on Local Stereopsis and Neuronal Responses of V1 and V2 in Prism-Reared Monkeys," *J Neurophysiol*, vol. **97**, no. **4** 2612-26.
7. Astle, A. T., P. V. McGraw, and B. S. Webb. (2011) "Recovery of Stereo Acuity in Adults with Amblyopia," *BMJ Case Rep*, vol. **2011**.
8. Ding, J., and D. M. Levi. (2011) "Recovery of Stereopsis through Perceptual Learning in Human Adults with Abnormal Binocular Vision," *Proc Natl Acad Sci U S A*, vol. **108**, no. **37** E733-41.
9. Xi, J., W. L. Jia, L. X. Feng, Z. L. Lu, and C. B. Huang. (2014) "Perceptual Learning Improves Stereoacuity in Amblyopia," *Invest Ophthalmol Vis Sci*, vol. **55**, no. **4** 2384-91.
10. Knill, D. C. (2005) "Reaching for Visual Cues to Depth: The Brain Combines Depth Cues Differently for Motor Control and Perception," *J Vis*, vol. **5**, no. **2** 103-15.
11. Knill, D. C., J. A. Saunders (2003) "Do humans optimally integrate stereo and texture information for judgments of slant" *Vis Res*, vol. **43**, 2539-58.
12. Simons, K. (1981) "A Comparison of the Frisby, Random-Dot E, Tno, and Randot Circles Stereotests in Screening and Office Use," *Arch Ophthalmol*, vol. **99**, no. **3** 446-52.
13. McKee, S. P., and D. M. Levi. (1987) "Dichoptic Hyperacuity: The Precision of Nonius Alignment," *J Opt Soc Am A*, vol. **4**, no. **6** 1104-8.
14. Acerbi, L., D.M. Wolpert, and S. Vijayakumar. (2012) Internal Representations of Temporal Statistics and Feedback Calibrate Motor-Sensory Interval Timing. *PLoS Comput Biol* **8**(11): e1002771. doi:10.1371/journal.pcbi.1002771
15. Levi, D. M., and R. W. Li. (2009) "Perceptual Learning as a Potential Treatment for Amblyopia: A Mini-Review," *Vision Res*, vol. **49**, no. **21** 2535-49.
16. Levi, D. M. (2012) "Prentice Award Lecture 2011: Removing the Brakes on Plasticity in the Amblyopic Brain," *Optom Vis Sci*, vol. **89**, no. **6** 827-38.
17. Hess, R. F., and B. Thompson. (2015) "Amblyopia and the Binocular Approach to Its Therapy," *Vision Res*, vol. **114** 4-16.
18. Tsirlin, I., L. Colpa, H. C. Goltz, and A. M. Wong. (2015) "Behavioral Training as New Treatment for Adult Amblyopia: A Meta-Analysis and Systematic Review," *Invest Ophthalmol Vis Sci*, vol. **56**, no. **6** 4061-75.
19. Ernst, M. O., and M. S. Banks. (2002) "Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion," *Nature*, vol. **415**, no. **6870** 429-33.
20. Ernst, M. O., M. S. Banks, and H. H. Bulthoff. (2000) "Touch Can Change Visual Slant Perception," *Nat Neurosci*, vol. **3**, no. **1** 69-73.
21. Knill, D. C., and J. A. Saunders. (2003) "Do Humans Optimally Integrate Stereo and Texture Information for Judgments of Surface Slant?" *Vision Res*, vol. **43**, no. **24** 2539-58.
22. Knill, D. C., and D. Kersten. (2004) "Visuomotor Sensitivity to Visual Information About Surface Orientation," *J Neurophysiol*, vol. **91**, no. **3** 1350-66.
23. Nardini, M., R. Bedford, and D. Mareschal. (2010) "Fusion of visual cues is not mandatory in children," *Proc Natl Acad Sci U S A*, vol. **107**, no. **39** 17041-46.
24. Ogle, K. N. (1952) "Disparity Limits of Stereopsis," *AMA Arch Ophthalmol*, vol. **48**, no. **1** 50-60.
25. Westheimer, G. And I. J. Tanzman. (1956) "Qualitative depth localization with diplopic images," *J Opt Soc Am*, vol. **46**, no. **2** 116-17.
26. Blakemore, C. (1970) "The Range and Scope of Binocular Depth Discrimination in Man," *J Physiol*, vol. **211** 599-622.
27. Dengler, B., and G. Kommerell. (1993) "Stereoscopic Cooperation between the Fovea of One Eye and the Periphery of the Other Eye at Large Disparities. Implications for Anomalous Retinal Correspondence in Strabismus," *Graefes Arch Clin Exp Ophthalmol*, vol. **231**, no. **4** 199-206.
28. Kaye, M. (1978) "Stereopsis without Binocular Correlation," *Vision Res*, vol. **18**, no. **8** 1013-22.
29. Wilcox, L. M., J. M. Harris, and S. P. McKee. (2007) "The Role of Binocular Stereopsis in Monoptic Depth Perception," *Vision Res*, vol. **47**, no. **18** 2367-77.
30. Foley, J. M., T. H. Applebaum, and W. A. Richards. (1975) "Stereopsis with Large Disparities: Discrimination and Depth Magnitude," *Vision Res*, vol. **15**, no. **3** 417-21.
31. Wilcox, L. M., and R. S. Allison. (2009) "Coarse-Fine Dichotomies in Human Stereopsis," *Vision Res*, vol. **49**, no. **22** 2653-65.
32. Giaschi, D., S. Narasimhan, A. Solski, E. Harrison, and L. M. Wilcox. (2013a) "On the Typical Development of Stereopsis: Fine and Coarse Processing," *Vision Res*, vol. **89** 65-71.
33. Giaschi, D., R. Lo, S. Narasimhan, C. Lyons, and L. M. Wilcox. (2013b) "Sparing of Coarse Stereopsis in Stereo-deficient Children with a History of Amblyopia," *J Vis*, vol. **13**, no. **10**.
34. Ding, J., S.A. Klein, and D.M. Levi. (2013) "Binocular Combination in Abnormal Binocular Vision," *J. Vis*.
35. Vedamurthy, I., M. Nahum, D. Bavelier, and D. M. Levi. (2015) "Mechanisms of Recovery of Visual Function in Adult Amblyopia through a Tailored Action Video Game," *Sci Rep*, vol. **5** 8482.
36. Welchman, A. E., A. Deubelius, V. Conrad, H. H. Bulthoff, and Z. Kourtzi. (2005) "3d Shape Perception from Combined Depth Cues in Human Visual Cortex," *Nat Neurosci*, vol. **8**, no. **6** 820-7.
37. Morgan, M. L., G. C. Deangelis, and D. E. Angelaki. (2008) "Multisensory Integration in Macaque Visual Cortex Depends on Cue Reliability," *Neuron*, vol. **59**, no. **4** 662-73.
38. Ban, H., T. J. Preston, A. Meeson, and A. E. Welchman. (2012) "The Integration of Motion and Disparity Cues to Depth in Dorsal Visual Cortex," *Nat Neurosci*, vol. **15**, no. **4** 636-43.
39. Murphy, A. P., H. Ban, and A. E. Welchman. (2013) "Integration of Texture and Disparity Cues to Surface Slant in Dorsal Visual Cortex," *J Neurophysiol*, vol. **110**, no. **1** 190-203.
40. Li, R. W., C. V. Ngo, and D. M. Levi. (2015) "Relieving the Attentional Blink in the Amblyopic Brain with Video Games," *Sci Rep*, vol. **5** 8483.
41. Scarfe, P., and A. Glennerster. (2015) "Using High-Fidelity Virtual Reality to Study Perception in Freely Moving Observers," *J Vis*, vol. **15**, no. **9** 3.

Table 1

Type	Age (yrs)/ Gender	NDE	Refractive Correction	ETDRS VA (logMAR)	Ocular Alignment (prims diopters)		Stereo acuity (arc sec) Pre/Post	Treatment history
					Distance	Near		
<b>A1</b>	19/F	R	R: -12.00 DS/-2.00 DCx10 L: +0.50 DS/-2.75 DC x 5	R: 0.2 L: 0.0	Ortho	Ortho	200/20	Detected at age 4, no treatment
<b>A2</b>	19/M	L	R: plano L: +4.50 DS/ -1.00DC x 180	R: -0.1 L: 0.16	Ortho	3 XP	200/70	Detected at age 11, patched at age 11 for one year
<b>M1</b>	23/F	R	R: -5.00DS/-2.25DCx20 L: -6.50 DS/-2.25 DCx170	R: 0.02 L: 0.0	RXT AXT, 12	RXT AXT, 10	70/20	Detected at age 2, no treatment
<b>M2</b>	38/F	L	R: plano L: +1.75 DS/-0.50 DC x 166	R: -0.1 L: 0.06	12 LXT	12 LXT	>400/>400	Detected in infancy. Two surgeries to correct strabismus
<b>M3</b>	20/F	L	R: plano L: +2.25 DS/ -0.75 DC x 165	R: -0.1 L: 0.08	10 LXT, 2 L HypoT	10 LXT, 2 L HypoT	200/>400	Detected at age 2, patched for 2 yrs.
<b>M4</b>	30/F	R	R: -5.25 DS/ -2.00 DC x 2 L: -3.25 DS/ -0.50 DS x 15	R: 0.18 L: 0.14	4 RET	4RET	200/70	Detected at age 19, no treatment
<b>M5</b>	23/F	R	R: +3.50 DS/-0.5 DCx10 L: +0.75 DS	R: 0.16 L: -0.1	5 RET	5 RET	>400/70	Detected at age 13
<b>S1</b>	56/F	R	R: -1.00 DS/-1.00 DCx90 L: -1.00 DS/ -1.25 DC x 85	R: 0.04 L: 0.02	RXT AXT, 22	RXT AXT, 24	200/20	Detected at age < 10 yrs, no treatment
<b>S2</b>	55/F	R	R: +3.00 DS L: +3.75 DS/ +0.25 DC x 26	R: 0.12 L: 0.04	3RET	3 RET	>400/>400	Detected at age 3, patched few weeks
<b>S3</b>	52/F	R	R: -2.25 DS/ -0.75 DC x 30 L: -2.50 DS/ -0.75 DC x 170	R: 0.24 L: -0.06	6 RET, 2 LHyper T	6 RET, 2 LHyper T	400/400	Detected at age 2, patched for 1 yr
<b>S4</b>	47/F	R	R: -4.25 DS L: -4.25 DS	R: 0.20 L: 0.0	6 RET	5 RET	200/200	Detected at age 31, no treatment

Abbreviations: (1) Type; A, anisometropia; M, both strabismus and anisometropia; S, pure strabismus (2) NDE, non-dominant eye (R – right; L – left); (3) Ocular Alignment; Ortho, Orthophoria; XP, exophoria; XT, exotropia; ET, esotropia; AXT, Alternating exotropia; HyperT, hypertropia; HypoT, hypotropia; (4) Stereoacuity; Randot stereoacuity. F, failed (>400 arcsec); Note that treatment history includes any treatment beyond refractive correction with glasses or contact lenses. Age appropriate near correction was used for the various test distances. Units: visual acuity is given in logMAR units. Subjects A1, A2, M1, M4 and M5 were contact lens wearers.

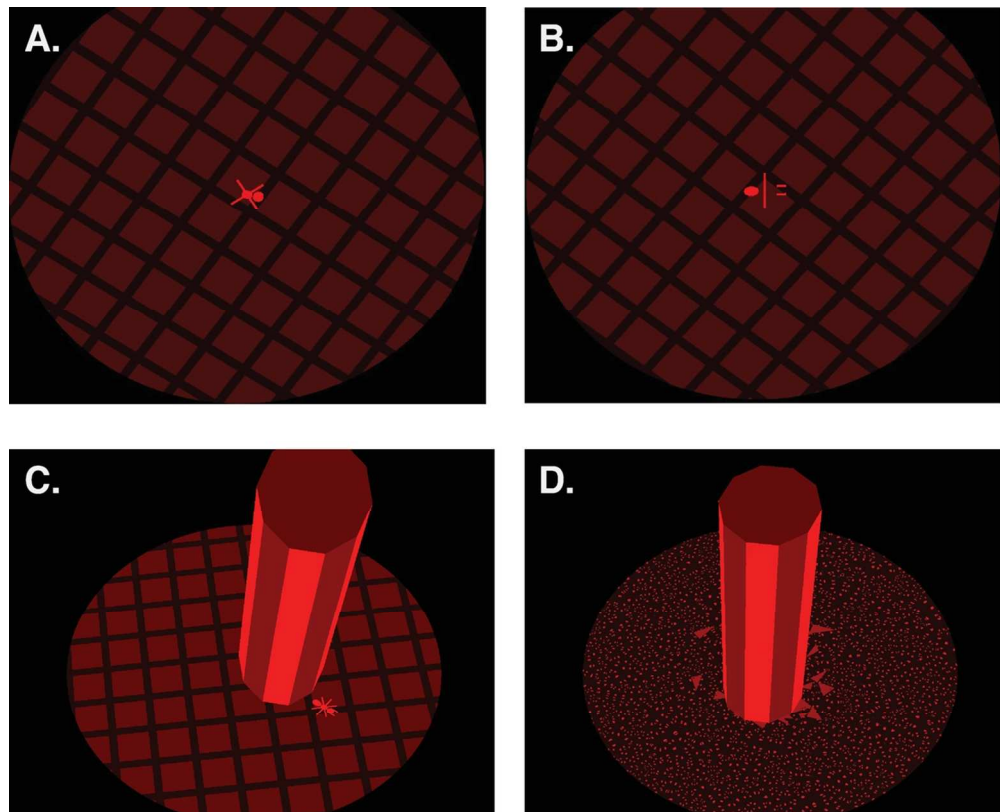


Figure 1. Examples of fixation 'bug' stimulus and texture surfaces. A) The fixation bug's thorax, abdomen and 4 diagonal legs were presented to the dominant eye. B) The head and remaining bug parts were presented to the non-dominant eye. When fused, a complete bug with six legs and a pair of antennae would be perceived. Note that Fig. 1A and 1B included here are for illustrative purposes only and were not designed as a stereo pair.

C) Tiled texture background, (D) Random dot background.

For tiled backgrounds rendered stereoscopically, the monocular cues to surface slant stem from distortions of the regular grid pattern mapped onto the surface, and the disparity signal to surface slant stems from the differences between right and left eyes' retinal images. Examples of trial feedback are also depicted here. The dichoptic bug crawls away (C) for an incorrect response, and it explodes in little bits (D) for a correct response.

107x86mm (300 x 300 DPI)

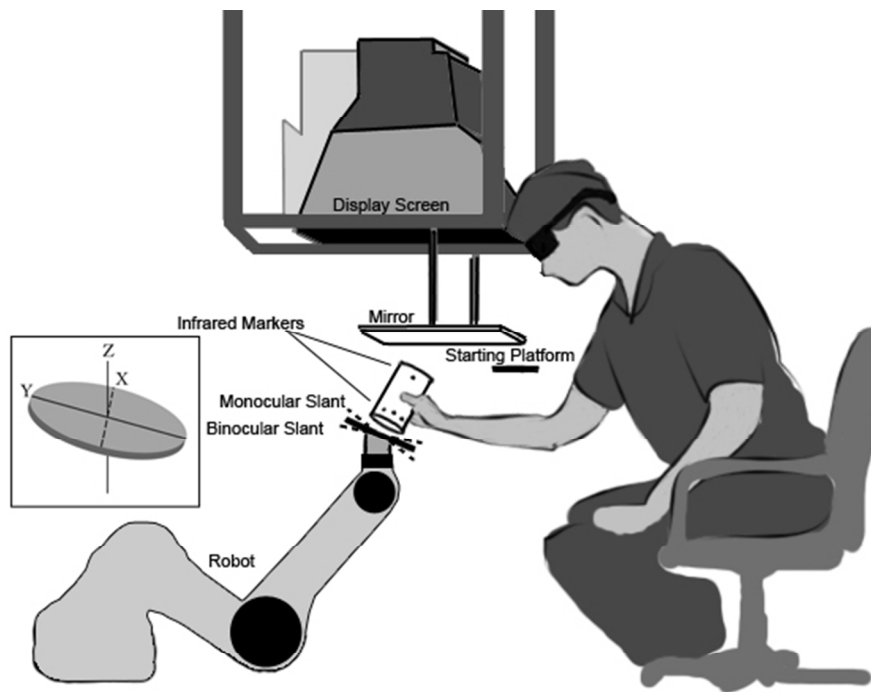
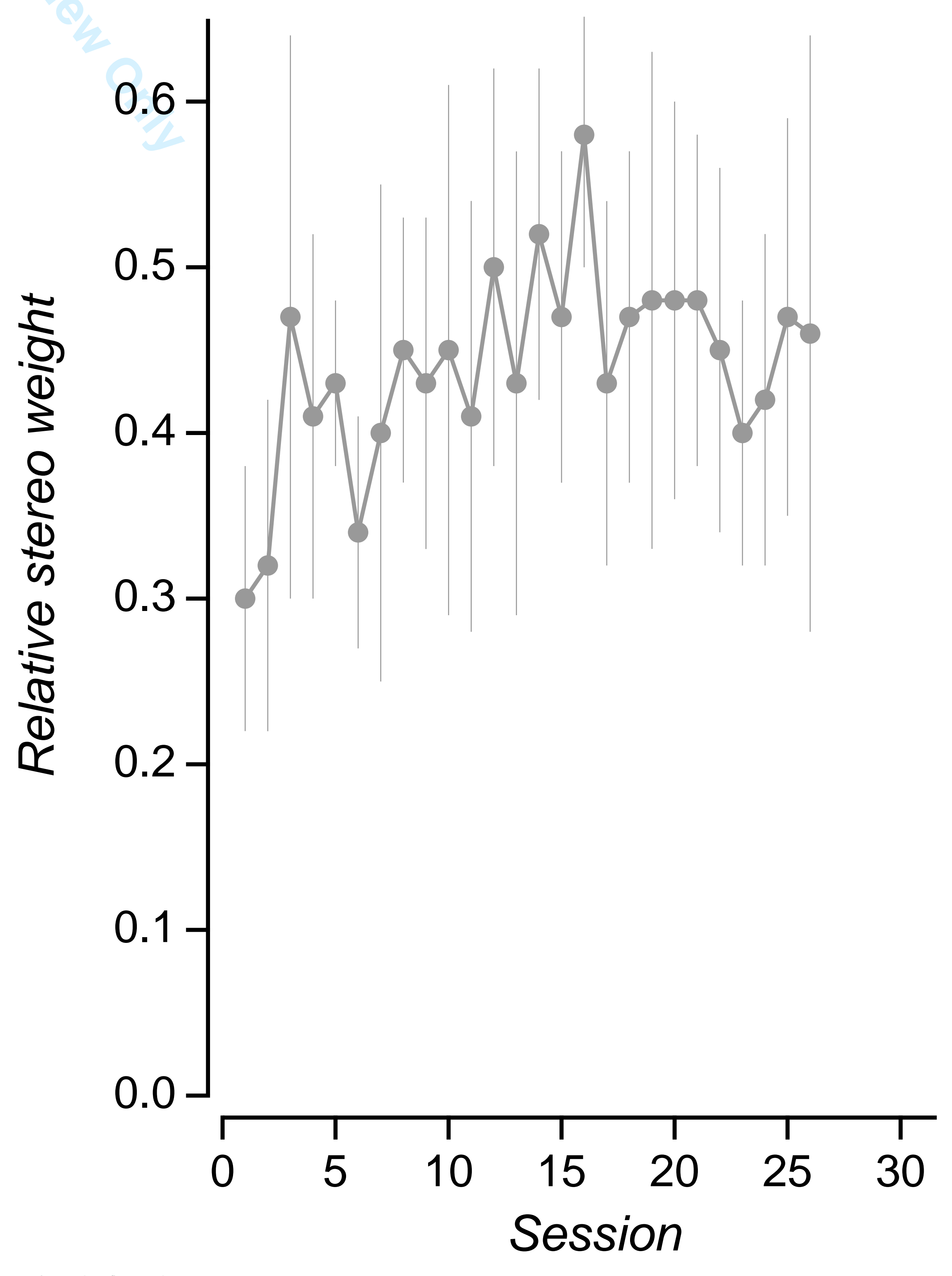
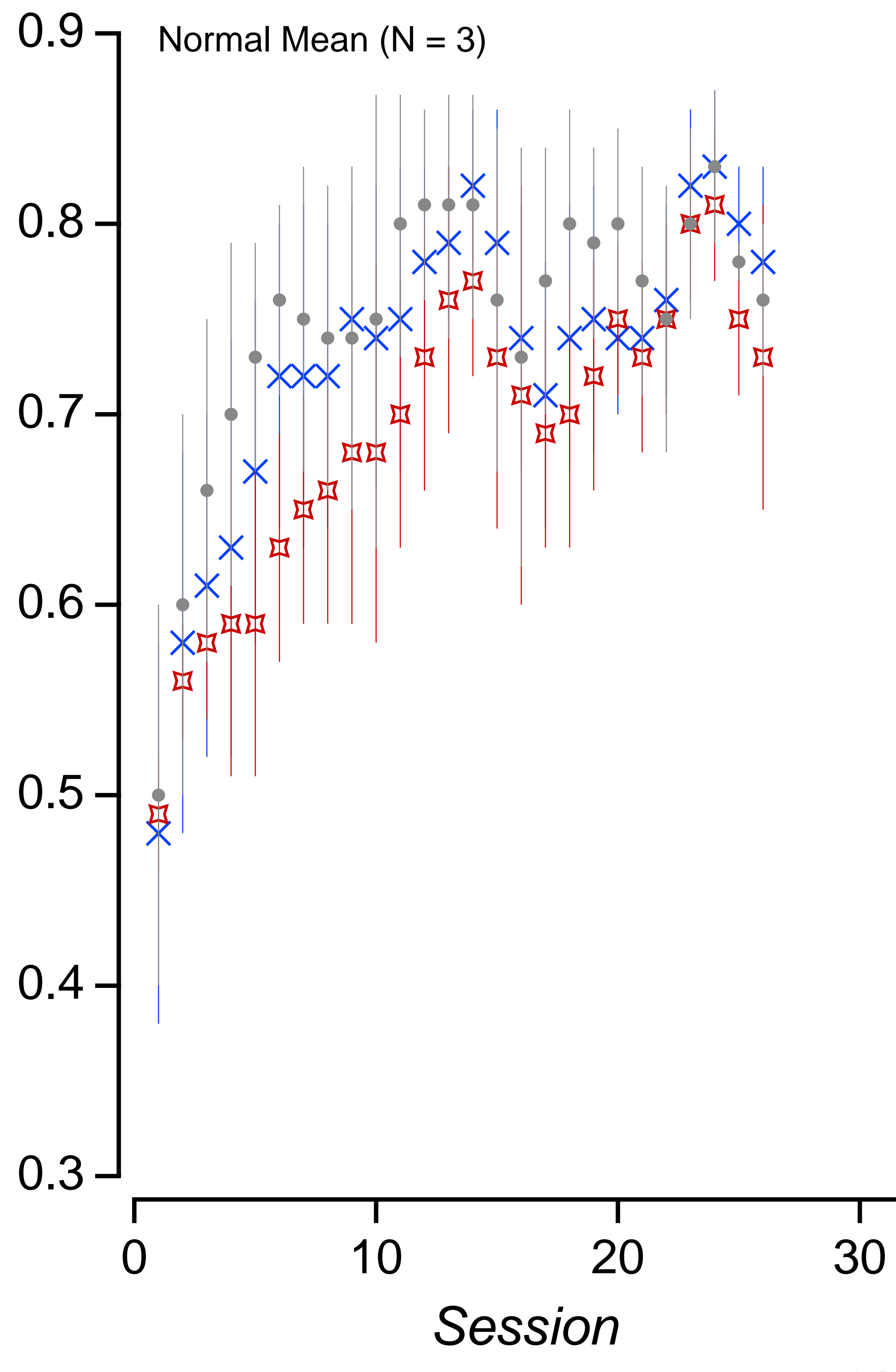
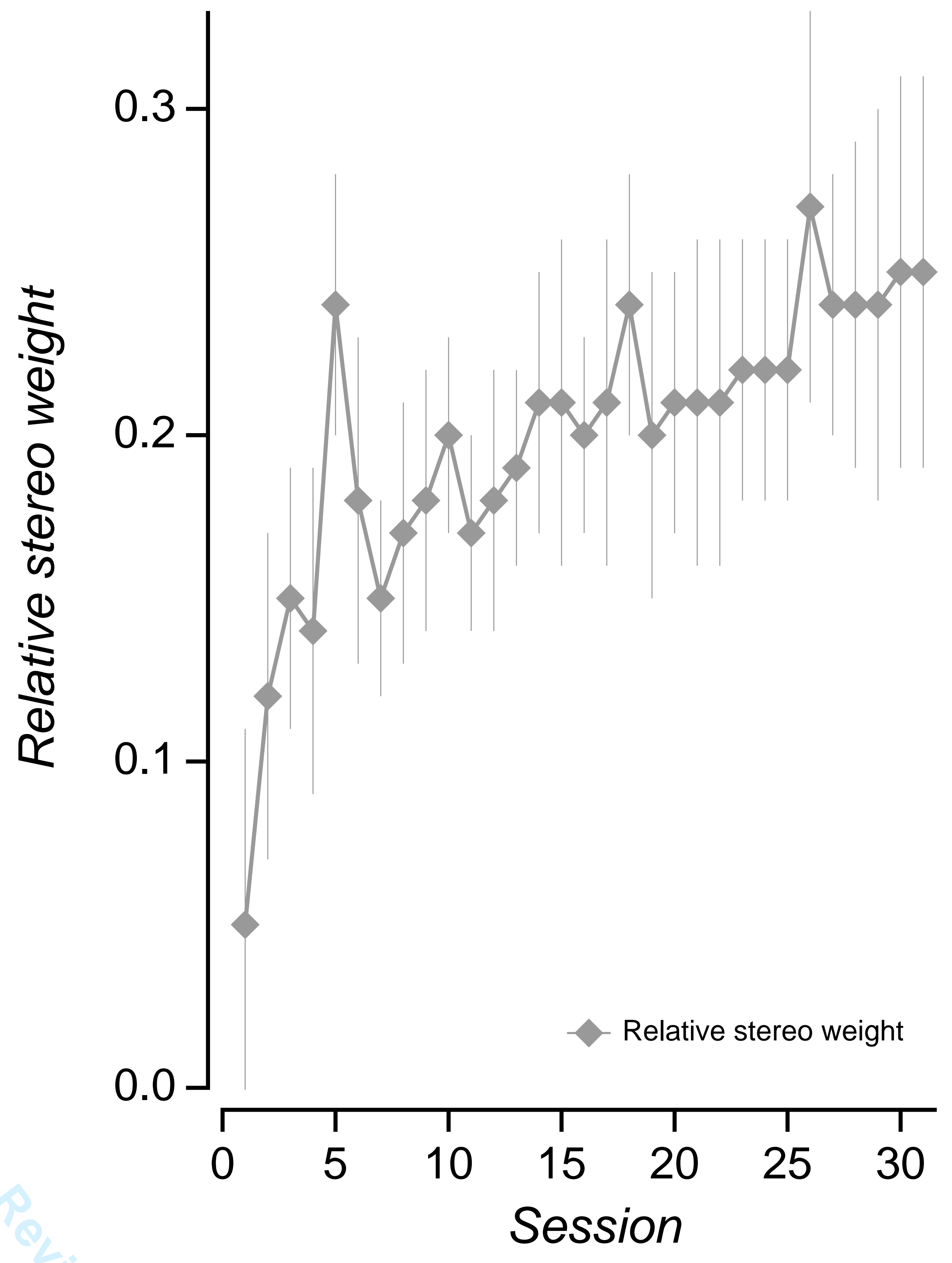
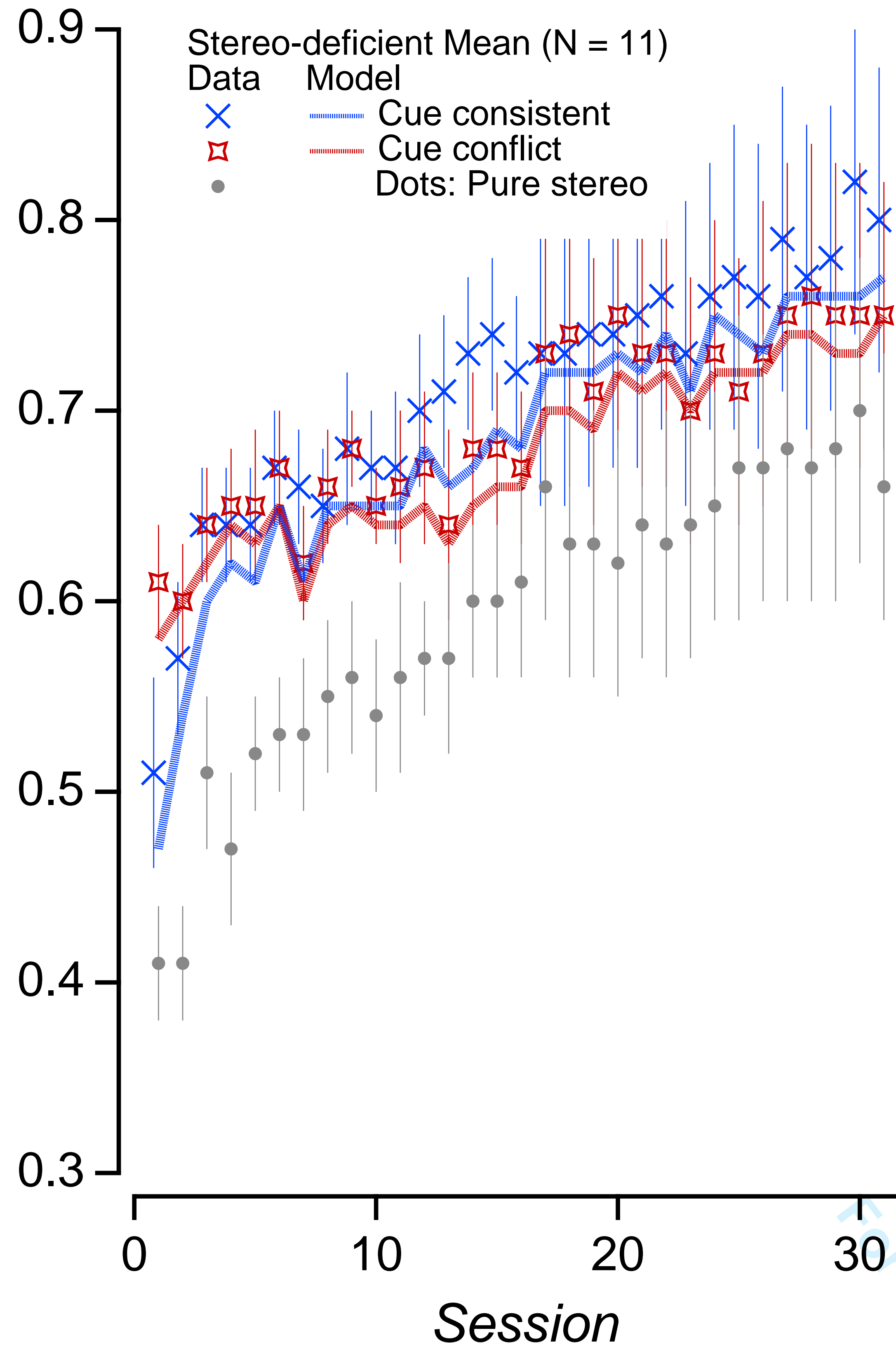


Figure 2. Stimuli were displayed on an inverted 22-inch Mitsubishi Diamond Pro 2070SB CRT monitor, with screen resolution 1152 x 864 pixels and refresh rate 120 Hz. Observers viewed the reflections of the stereoscopically presented stimuli through a mirror using Crystal Eyes shutter glasses (StereoGraphics Corporation, San Rafael, CA). An opaque black plate was placed beneath the mirror so that subjects could see only the virtual image of the display formed below the mirror. A PUMA 260 robot arm, invisible to the subject, coaligned a circular metal (target) surface with the virtual image at a distance of 63.5cm from the viewer's eyes. The slant of the target surface was defined as its orientation around the x-axis relative to subject's line of sight (see dashed line in the figure inset), and a zero degree slant indicated frontoparallel. Monocular and binocular slants differed on cue-conflict trials, but subjects perceived the stimuli as a single slanted disk. Head movements were restricted with head and chin rests. Subjects' task was to move a cylinder from the starting platform and place it flush onto the real surface to squash the virtual image. The slant of the real surface was determined by the mean of the monocular and binocular slants. The starting platform was 40 cm to the right of the target surface, 20 cm closer to the subject than the target surface, and 16.5 cm above the target surface (all measured from the subject's point of view).  
223x167mm (72 x 72 DPI)

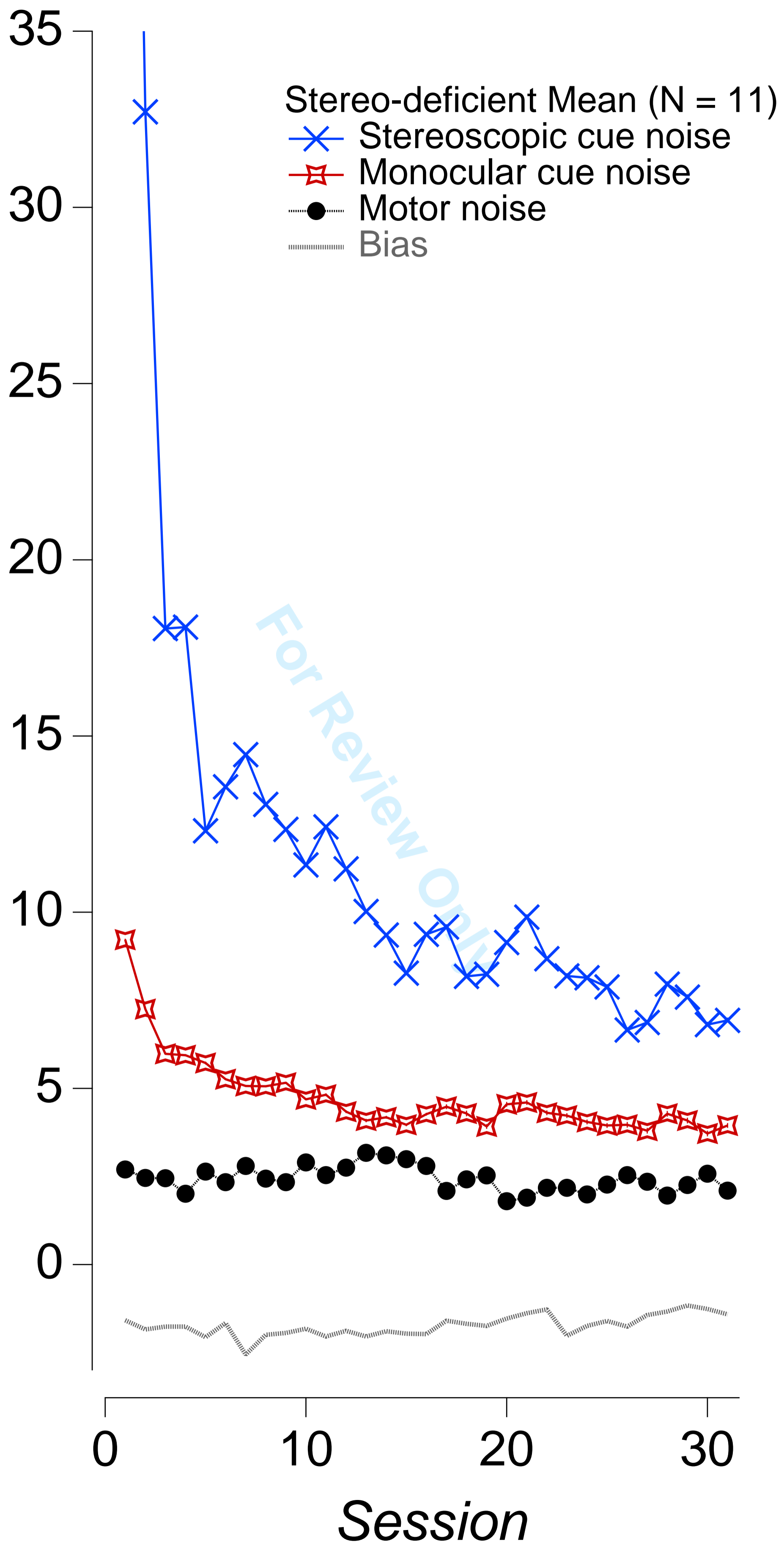


1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

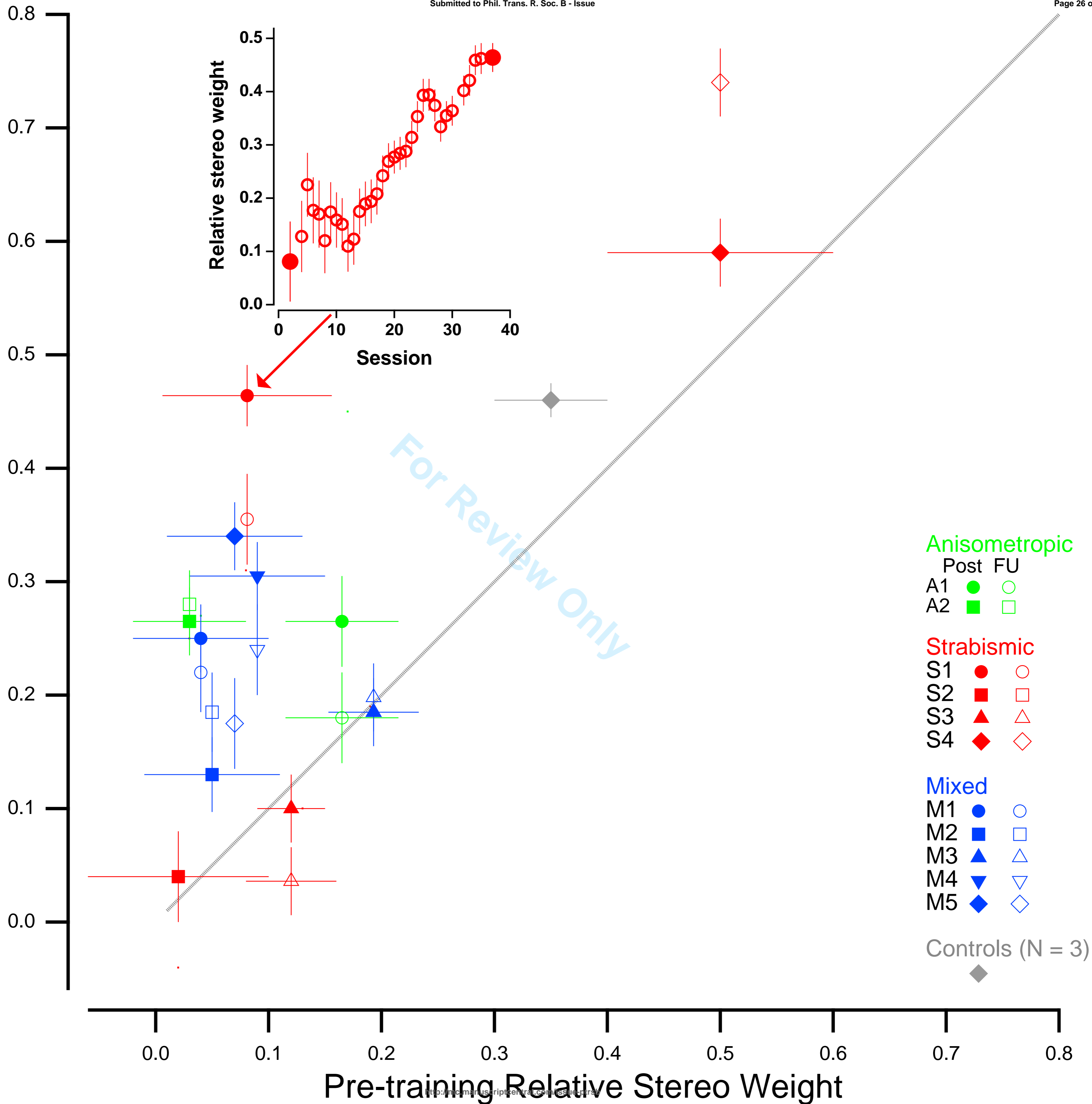




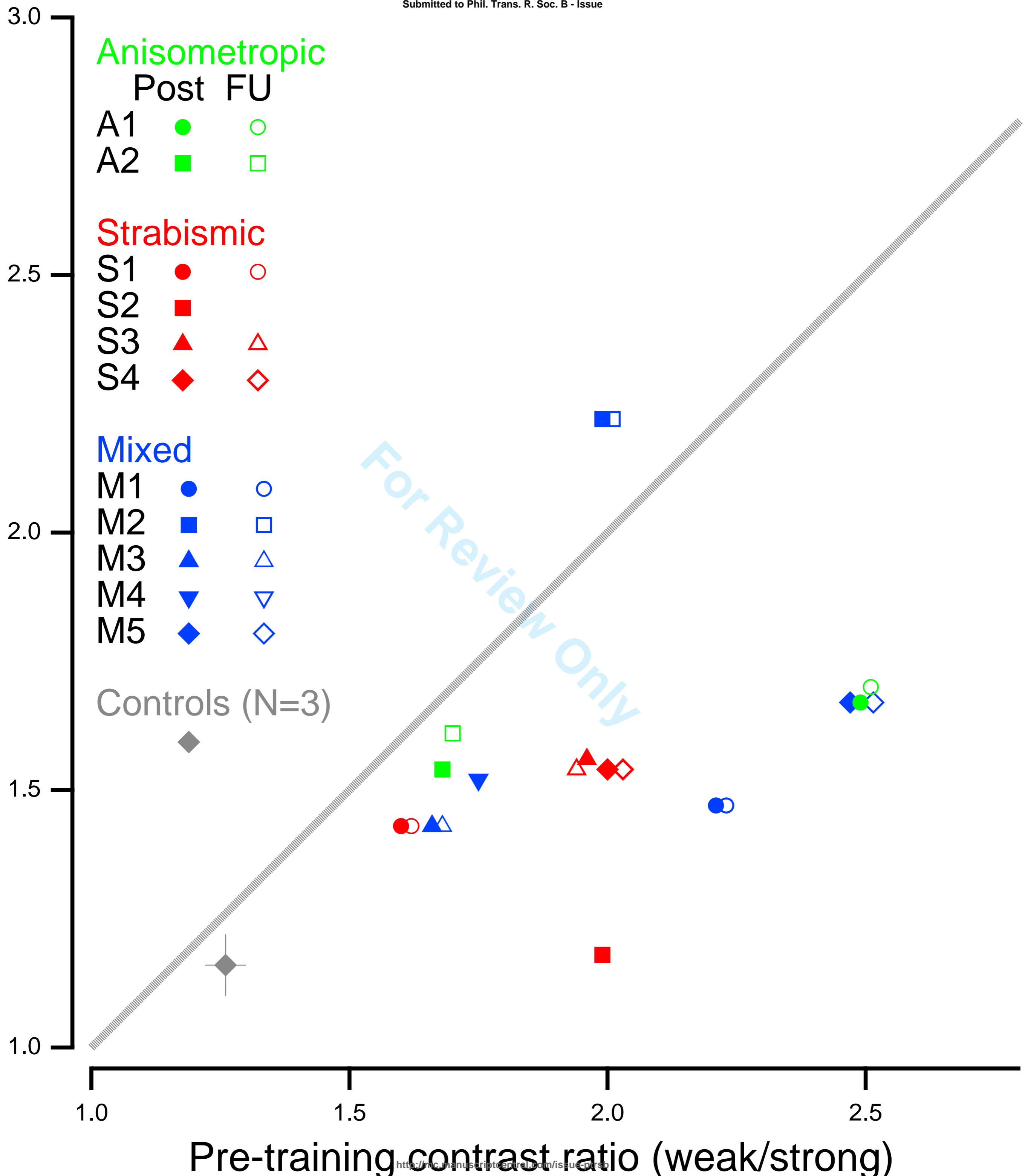
Sensory noise S.D. (degrees)

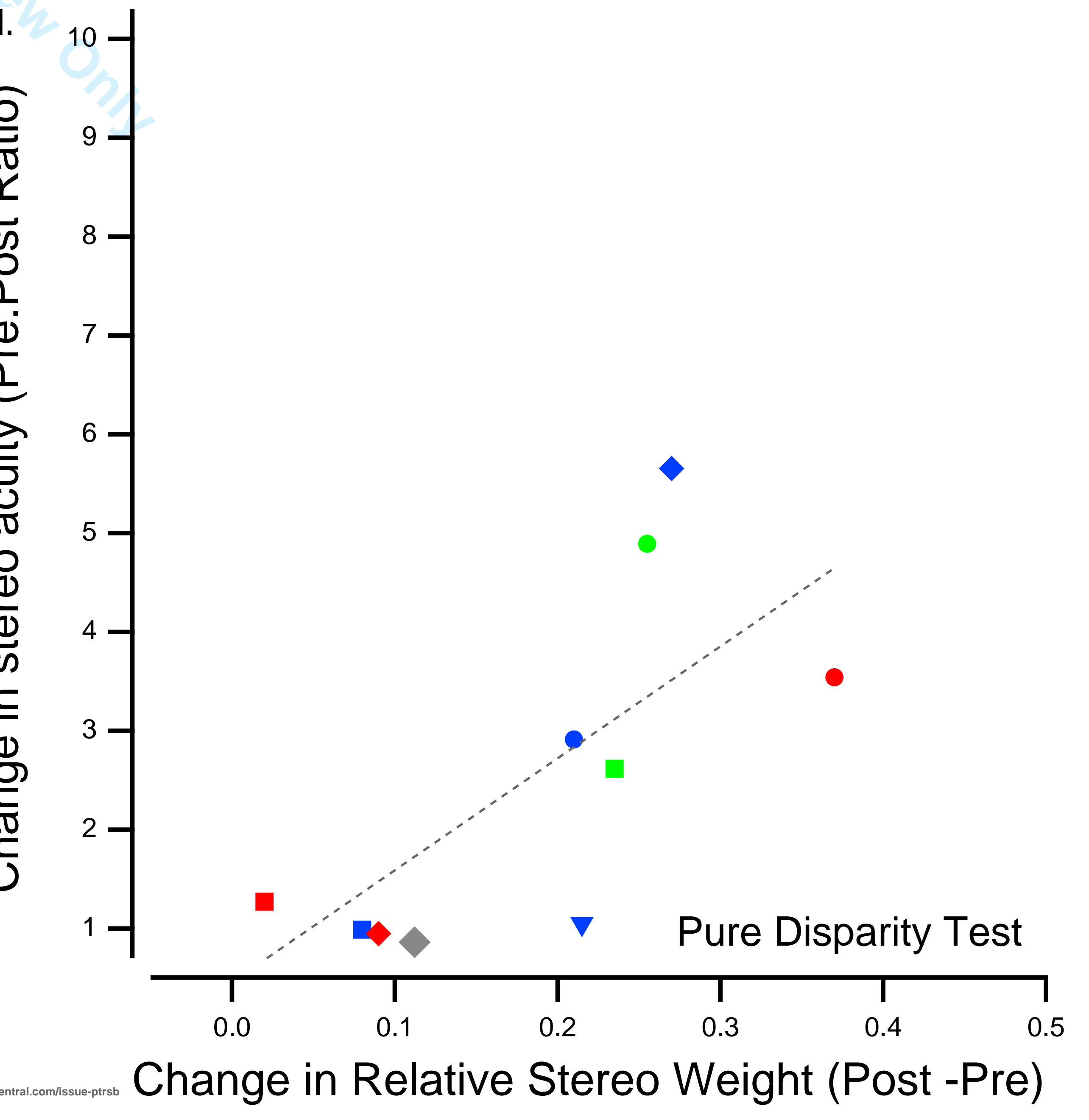
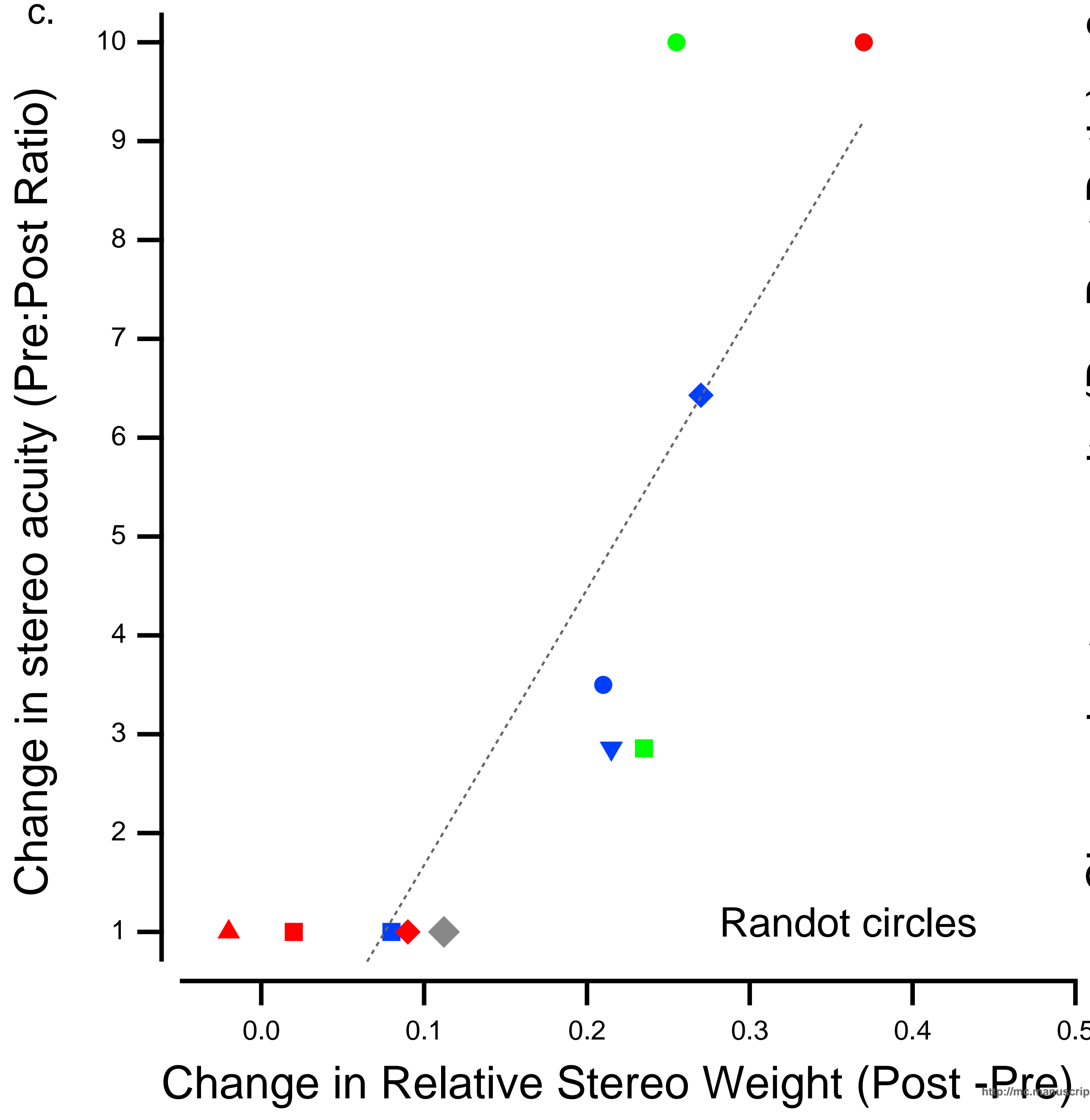
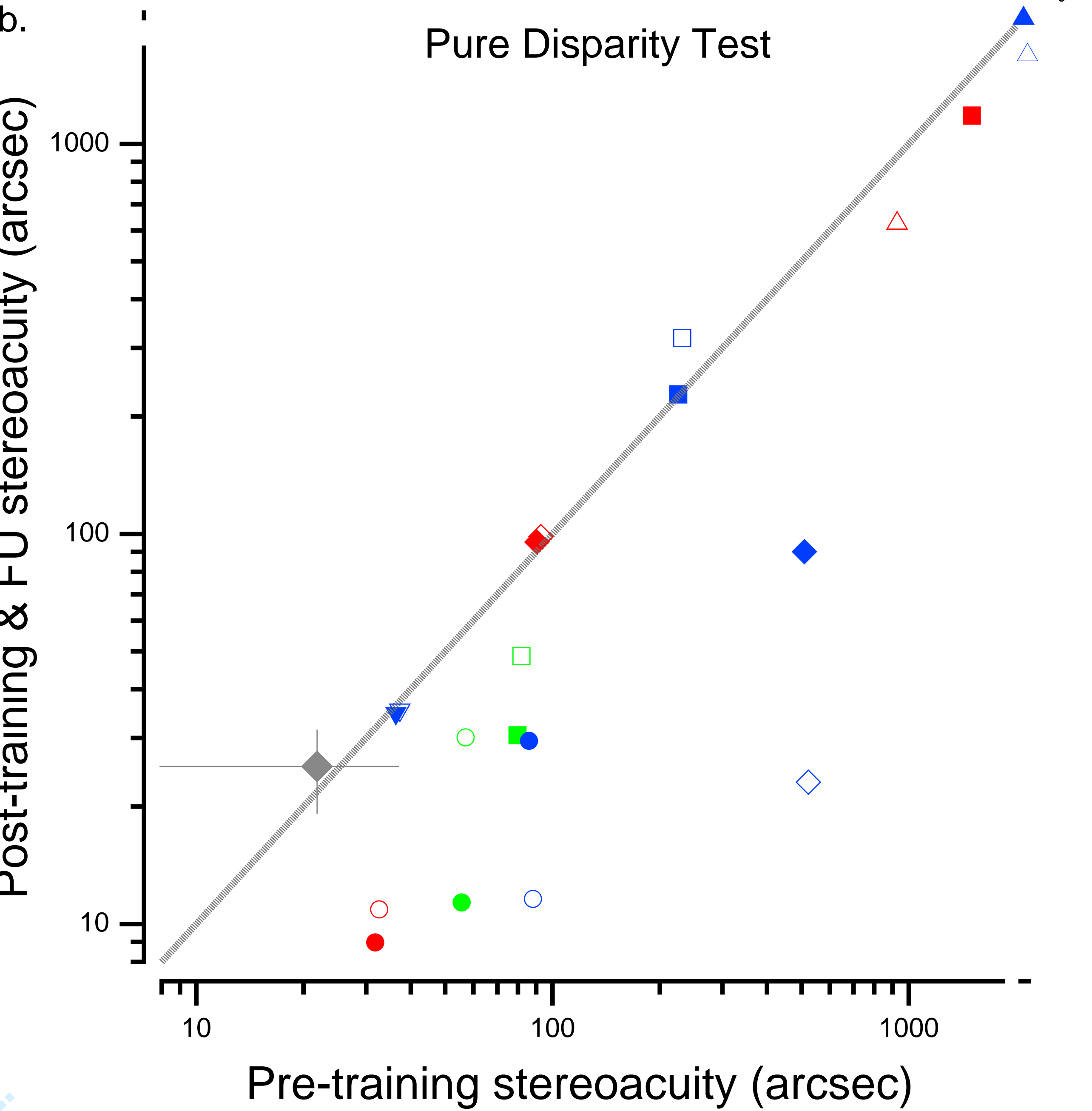
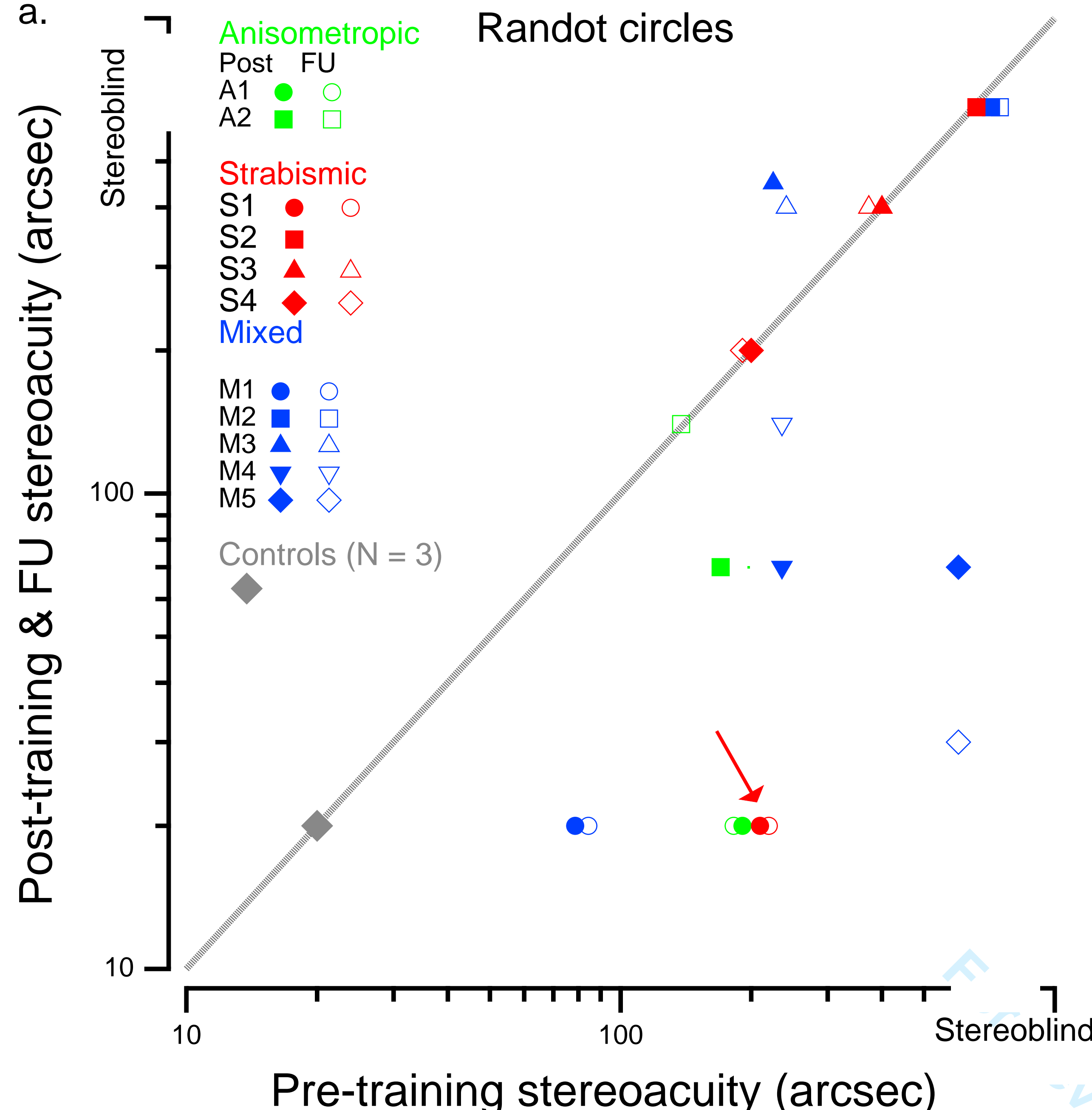


## Post-training &amp; FU Relative Stereo Weight

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60





1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### Anisometric

Post FU

- A1 ● ○
- A2 ■ □

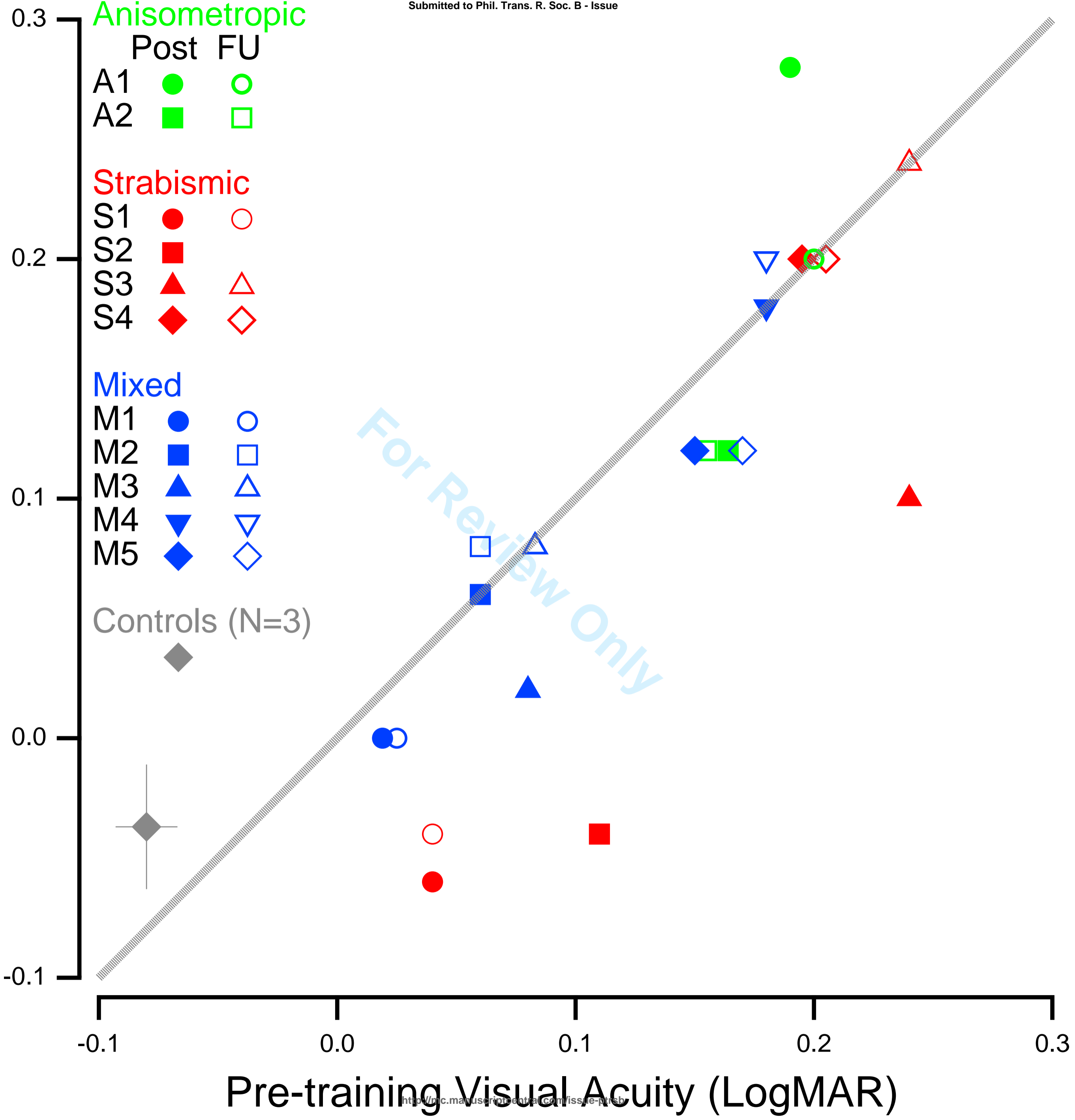
### Strabismic

- S1 ● ○
- S2 ■ □
- S3 ▲ △
- S4 ◆ ◇

### Mixed

- M1 ● ○
- M2 ■ □
- M3 ▲ △
- M4 ▼ ▽
- M5 ◆ ◇

Controls (N=3)



For Review Only

Pre-training Visual Acuity (LogMAR)