**Method.**

*Participants.* Twenty adults (mean age: 28 ± 2.5, range: 18 – 62 years) with normal or corrected-to-normal vision and no ocular pathologies participated the training study. All participants completed 40 hours of training with the exception of DL and AJ who completed 20 hours total. Participants were recruited by telephone from the Meredith W. Morgan University Eye Center’s internal list of patients who gave written consent to be contacted for research studies and through internal UC Berkeley email list serves. The study protocol was approved by the Institutional Review Board of the University of California, Berkeley.

Exclusion criteria for the study included: (1) ocular pathologies (e.g., macular abnormalities) or nystagmus, (2) non-concomitant or large angle constant strabismus (> 30 prism diopter), (3) inability to fuse, (4) constant esotropia (> 20 prism diopters), (5) visual acuity (VA) ≥ 20/200, and (6) dichoptic visual training of more than 10 hours.

*Screening.* All participants were screened by a UC Berkeley Optometry student before and after the study. Measurements collected include: (1) visual acuity, (2) current prescription, (3) refraction at distance, (4) clinical stereoacuity (Randot), (5) ocular deviation (cover test), (6) binocular fusion (Worth Dot), and (7) retinal health. Participants were categorized as anisometropic if there was a difference ≥ 0.50D in spherical equivalent refraction or ≥ 1.5D difference in astigmatism in any meridian, between the two eyes (Wallace et al., 2011). Strabismus was classified by a presence of tropia with the cover test. Clinical details of each participants can be found in Table 1.

*Study design and training.* Following consent and screening, eligible participants were placed in either the stereo-normal or stereo-anomalous group based on their initial Randot stereoacuity measurement. Inclusion criteria for the stereo-anomalous group included a clinical baseline stereoacuity worse than 40 arc secs. Once placed in a group, participants completed a series of tests to measure contrast sensitivity and stereoacuity. For contrast sensitivity, we used the quick CSF method (Hou et al., 2016) displayed on a 46” NEC LCD monitor (model p463) with a resolution of 1920 x 1080 and a contrast ratio of 4,000:1.

To measure, we administered both clinical and psychophysical tests. For a clinical stereoacuity measurement, we used the clinical random-dot stereogram Randot circles and Randot 3 with polarized glasses. For psychophysical tests, we used the Pure Disparity Test (PDT) and dynamic random-dot stereogram (DRS) (Ding & Levi, 2011). Both PDT and DRS stimuli were viewed through a stereoscope and presented on a Sony CRT monitor (need model) at a viewing distance of 60 cm from the participant.

Following the pre-tests, participants alternated every 10 hours between the Halloween and DartBoard game until they reached 40 hours of training (except for DL and AJ who completed 20 hours). After every 10 hours, the tests were administered to monitor learning. At the end of 40 hours, participants completed the screening that was administered by a UC Berkeley optometry student.

*Apparatus.* The games were presented on the Oculus Rift® Developer Kit (DK) 2, which is equipped with a gyroscope, accelerometer, and a magnetometer with an update rate of 1000 Hz. The Oculus Rift DK-2 has a resolution of 960 x 1080 for each eye, a field of view of 100 degrees, a refresh rate of 60-75 Hz, and a position tracking refresh rate of 60 Hz. To run the software, we used the Alienware AREA51R2 computer with Intel® Core™ i7-5820K CPU and an NVIDIA GeForce GTX 980 graphics card.

*Games.* We designed two video games (Halloween and DartBoard) to train stereovision using the Oculus Rift® VR glasses. Both games started with a synoptophoric high contrast dichoptic fusion-lock frame calibration (Fig 1), where deviation angle for strabismic participants could be corrected and suppression can be suspended by adjusting a virtual luminance filter. To correct for deviation angles, the researcher manually adjusted the images presented to each eye (horizontal, vertical, and cyclo deviations, plus aniseikonia) until the dichoptic cross presented at the center of the frame reached complete alignment for the participant. Thus, this step allowed the researcher to correct for any ocular deviation in subjective angle of squint. To avoid suppression, image luminance (ranging from equal luminance between both images to complete occlusion of one eye) was adjusted for the dominant eye until the participant perceived equal luminance of the dichoptic lines crossing the reference frame.

Both games were presented in three consecutive difficulty levels (Natural, Advanced, and Expert). Each session began with the Natural level, which consisted of monocular and binocular cues to depth such as shadows, perspective, motion parallax, and binocular disparity. In the Advanced level, shadows were no longer available and perspective was reduced, which meant that relative size was no longer a cue (i.e., object size was not relative to object distance and varied from trial to trial). Lastly, in the Expert level, motion parallax is limited (only rotational movements of the head were allowed by the software), making binocular disparity clearly the most helpful (almost unique) cue to calculate distance. Anti-suppression tasks by means of dichoptic images were also inserted in each games’ mechanics.

The difference between Halloween and DartBoard, apart from their game mechanics, lies in the stereoscopic strategy best suited for each stimulus. In DartBoard, the user is instructed to launch a dart, which is presented at the center of the screen, towards a dartboard that comes from the back of the scene and moves in the z-direction towards the front of the scene. Movement of the dart is always linear and travels from left to right, while movement of the board is linear and travels from back to front (and not necessarily in a 90 degrees angle): both linear movements occur in the same plane where the observers’ eyes are, avoiding monocular clues through intersection guessing. The user is then tasked with estimating the distance of the dart and the distance of the moving board independently, using the background wall as reference. Thus, performance in the DartBoard game can be interpreted as a measure of absolute disparity. In Halloween, the participant is instructed to shoot the closest target in a variable set (three to seven) of tragets as they approach the observer. Thus, Halloween is a conventional relative depth perception task, where the observer is tasked with estimating relative depth between targets.