

# Factors limiting sensitivity to binocular disparity in human vision: Evidence from a noise-masking approach

**Sara Alarcon Carrillo**

McGill Vision Research, Department of Ophthalmology  
& Visual Sciences, McGill University,  
Montreal, Quebec, Canada



**Alex S. Baldwin**

McGill Vision Research, Department of Ophthalmology  
& Visual Sciences, McGill University,  
Montreal, Quebec, Canada



**Robert F. Hess**

McGill Vision Research, Department of Ophthalmology  
& Visual Sciences, McGill University,  
Montreal, Quebec, Canada



Our visual system uses the disparity between the images received by the two eyes to judge three-dimensional distance to surfaces. We can measure this ability by having subjects discriminate the disparity of rendered surfaces. We wanted to know the basis of the individual differences in this ability. We tested 53 adults with normal vision using a relative disparity detection task. Targets were wedge-shaped surfaces formed from random dots. These were presented in either crossed or uncrossed disparity relative to a random dot background. The threshold disparity ranged from 24 arc seconds in the most-able subject to 275 arc seconds in the least-able subject. There was a small advantage for detecting crossed-disparity targets. We used the noise-masking paradigm to partition subject performance into two factors. These were the subject's equivalent internal noise and their processing efficiency. The parameters were estimated by fitting the linear amplifier model. We found both factors contributed to the individual differences in stereoacuity. Within subjects, those showing an advantage for one disparity direction had enhanced efficiency for that direction. Some subjects had a higher equivalent internal noise for one direction that was balanced out by an increased efficiency. Our approach provides a more thorough account of the stereo-ability of our subjects compared with measuring thresholds alone. We present a normative set of results that can be compared with clinical populations.

## Introduction

The human visual system can make depth judgements using binocular disparity (Wheatstone, 1838). Convergence of the eyes establishes a horopter. Any point on this imaginary surface maps onto corresponding locations on the retina of each eye (Schor, 2013). Objects closer to the subject than the horopter map onto locations on the retinae shifted “outward” (temporally). Those further than the horopter map onto inward (nasally) shifted locations. Shifts in these two directions result in crossed or uncrossed disparity, respectively. Evidence from animal neurophysiology indicates that the disparity of a stimulus is encoded by binocular cells in the visual cortex (Barlow, Blakemore, & Pettigrew, 1967; DeAngelis, Ohzawa, & Freeman, 1991; Poggio & Fischer, 1977). Behaviorally, humans are very sensitive to binocular disparity (McKee, Welch, Taylor, & Bowne, 1990). Human stereoacuity thresholds smaller than 5 seconds of arc (arcsec) have been obtained (McKee, 1983; Stevenson, Cormack, & Schor, 1989). Previous studies, however, have shown considerable individual differences in stereoacuity across the population (Bosten et al., 2015; Dorman & van Ee, 2017; Hess et al., 2015; Hess et al., 2016; Zaroff, Knutelska, & Frumkes, 2003). In this study, we investigate the variability for a relative disparity detection task. We also compare performance for two different directions of depth relative to the stimulus background. We then use noise-masking to explore the individual differences we found.

Citation: Alarcon Carrillo, S., Baldwin, A. S., & Hess, R. F. (2020). Factors limiting sensitivity to binocular disparity in human vision: Evidence from a noise-masking approach. *Journal of Vision*, 20(3):9, 1–14, <https://doi.org/10.1167/jov.20.3.9>.



Several methods can be used to assess stereoacuity. Individual differences in a population's measured stereoacuity depend on the task used (Bohr & Read, 2013; Coutant & Westheimer, 1993; Newhouse & Uttal, 1982; Patterson & Fox, 1984). One basic task is having the subject detect a difference in disparity in a stimulus (i.e., a measure of relative disparity sensitivity, which we term a relative disparity detection task here). A previous study using this method (Bosten et al., 2015) tested the ability to detect relative crossed and uncrossed disparity. Approximately 10% of their subjects failed to detect either crossed or uncrossed disparity (thresholds greater than 350 arcsec). Approximately 5% of the subjects failed the task for both directions. Approximately 3% of the population in that study had amblyopia. Discounting these subjects, at least 12% of the nonamblyopic subjects had abnormally high thresholds for one or both disparity directions. The term “stereoanomalous” refers to subjects with unexplained stereo impairments of this type.

Higher incidences of stereoanomaly have been reported by studies using a depth-polarity discrimination task. When the direction of disparity must be discriminated (crossed vs. uncrossed), the proportion of stereoanomalous individuals increases to approximately 30%. This was first shown in a task in which targets were flashed in isolation (R. Jones, 1977; Richards, 1970; Richards, 1971). It has since been demonstrated in a task in which subjects judge the relative disparity of a target compared with a reference plane (Hess et al., 2015). Further evidence has been provided from a task in which subjects must make judgements about the range of depths in a “cloud” of rendered bars (van Ee & Richards, 2002). An explanation has been put forward based on separate mechanisms tuned to crossed or uncrossed disparity. Previous studies suggest that the visual system is more sensitive to crossed disparity (Landers & Cormack, 1997; Manning, Finlay, Neill, & Frost, 1987; Patterson et al., 1995). Sensitivity to crossed disparity also develops earlier in infancy (Birch, Gwiazda, & Held, 1982). Difficulties in polarity-based tasks could be due to a lack of mechanisms tuned to one depth polarity (Richards, 1971). Alternatively, these results can also be explained by a model in which stereo processing occurs in a series of stages. If the *direction* of the disparity is determined at a relatively late stage, then it may be this step that is failing (Dorman & van Ee, 2017; Landers & Cormack, 1997).

The disparity of a stimulus depends on the position of the horopter. Therefore the vergence of the eyes is a critical factor. Presentation of a stimulus in coarse disparity relative to fixation will typically elicit vergence eye movements. R. Jones (1977) investigated the relationship between these eye movements and the stereoanomaly reported by (Richards, 1970). In the

study by Richards, the target stimuli (vertical bars) were presented alone (the fixation mark disappeared during stimulus presentation). Thresholds for identifying the direction of disparity from such targets are relatively high. These tasks are said to test coarse stereopsis. R. Jones (1977) tested a population of 30 subjects. These had all achieved a “normal” stereoacuity score on a clinical test based on relative disparity. Therefore they were stated to have no impairment in their fine stereopsis ability. In replicating the Richards (1970) task, only 12 subjects behaved normally. Eighteen subjects exhibited abnormal psychophysical behavior. Of these, a minority (six) were shown to also have an anomaly with the vergence eye movements that would normally be initiated on presentation of the target. Although all subjects who had anomalous vergence movements in R. Jones (1977) also exhibited stereoanomaly, this was not the case in a follow-up study (Fredenburg & Harwerth, 2001). It has been proposed that vergence anomalies during development may lead to stereoanomaly in later life (Dorman & van Ee, 2017).

Vergence errors in human eye movements further complicate the relationship between disparities presented in a stimulus and the retinal locations they project to. The attempt to fixate the eyes on a point in depth results in both an offset error (convergence at the wrong depth) and a variability in convergence over time (Steinman, Cushman, & Martins, 1982; Ukwade, Bedell, & Harwerth, 2003). Under natural viewing conditions, offset errors of several minutes of arc (arcmin) are typical. In one study, errors of up to 120 arcmin were common (Cornell, Macdougall, Predebon, & Curthoys, 2003). For a stimulus presentation of two seconds, Krauskopf, Cornsweet, & Riggs (1960) found a vergence standard deviation of (at best) 30 arcsec. Without controlling for these vergence errors, it is not possible to assert that the reference plane in a stimulus was at zero disparity. Targets in crossed or uncrossed disparity relative to that reference plane may not be in absolute crossed or uncrossed disparity on the retina. This is especially true when targets are presented in finer disparities. Yet, previous behavioral studies (reviewed by Ukwade, Bedell, & Harwerth, 2003) have found that humans are surprisingly resilient to the effects of these vergence errors. They have relatively little impact on tasks in which subjects must make use of disparity information. The variability in eye vergence can still, however, obscure the effects of the crossed/uncrossed stereoanomaly proposed by Richards (1970). Subjects' eyes may converge in such a way that allows them to perform the task (as was demonstrated deliberately in the appendix of van Ee & Richards, 2002). In Hess et al. (2015), a stereoanomalous subpopulation is identified in a relative disparity task. Here the threshold disparities should be relatively small compared with the subjects' vergence errors.

In this study, we investigated the range of stereoacuity present in a population with normal vision. We measured stereoacuity using a relative disparity detection task. This was a modified version of a recently developed stereoacuity test (Tittes et al., 2019; Webber et al., 2018). The test presents random dot stereogram stimuli. These target disparity-sensitive processing mechanisms and avoid additional (nondisparity) cues (Julesz, 1960) that may allow subjects without stereo vision to solve the task (Serrano-Pedraza, Vancleef, & Read, 2016). We first measured sensitivity to targets in relative crossed and uncrossed disparity in 53 subjects with no visual pathologies. This gave us an estimate of the stereoacuity range in the normal population. We then selected 18 subjects for further investigation. We tested those subjects in a noise-masking experiment in which external disparity noise was added to the stimulus. This allows for analysis with the equivalent noise paradigm. This was originally applied to luminance (Cohn, 1976) and contrast detection (Pelli, 1981). Since then it has been applied to other visual modalities, including global form (Dakin, 2001), motion (Hess, Mansouri, Dakin, & Allen, 2006), contour integration (Baldwin, Fu, Farivar, & Hess, 2017), and stereoscopic disparity (Wardle, Bex, Cass, & Alais, 2012).

In noise-masking experiments, thresholds are typically unaffected by low levels of external noise. Once the external noise exceeds some critical value, however, thresholds increase linearly with its standard deviation. The transition point is the subject's "equivalent internal noise" for performing the task. When the external noise is much smaller than the equivalent internal noise then performance is limited by the internal noise. When the external noise is much greater than the internal noise then performance is limited by that external noise. A noise-masking approach can therefore be used to break down threshold sensitivity into these two factors (Pelli & Farell, 1999). The previous study applying this method to the detection of stereoscopic disparity sought to determine why sensitivity decreases at greater eccentricities from fixation. They found that equivalent internal noise increases with eccentricity, but processing efficiency remains constant (Wardle et al., 2012). Our study applies a similar paradigm to investigate differences in sensitivity between individuals. We also investigate differences *within* individuals in sensitivity to relative crossed and uncrossed disparity.

## Methods

### Participants

We tested 53 subjects (25 women, aged 20–67). Subjects 5, 33, and 50 are authors. Subjects reported

having vision that was normal or corrected-to-normal with prescribed optical correction. We did not include subjects reporting any visual disorders (cataracts, glaucoma, or amblyopia). For 40 of the subjects, we measured visual acuity with the Freiburg Vision Test (FrACT v3.9.9a; Bach, 1996). We used the four-alternative Tumbling E task at a distance of 160 cm, allowing for a best acuity measure of 20/10. Measured acuities ranged from 20/20 to 20/10. All subjects gave written informed consent. The experiments were performed in accordance with the Declaration of Helsinki and were approved by the Research Ethics Board of the McGill University Health Centre.

### Apparatus

Stimuli were created and presented on a Eurocom P370EM Scorpius laptop (Eurocom, Ottawa, Canada) with a 17.3-in. 120 Hz display. An NVIDIA 3D Vision 2 system (Santa Clara, California) was used for stereoscopic stimulus presentation. Stimuli were presented using frame-interleaving with synchronized shutter glasses. The monitor resolution was  $1920 \times 1080$  pixels with a mean luminance of  $197 \text{ cd/m}^2$ . The display was gamma-corrected with a VIEWPixx X-Rite iDisplay Pro (VPixx Technologies, Saint-Bruno, Canada). Viewing distance for the stereo tasks was 55 cm. The room was dark during testing with the test screen as the only light source.

### Stimulus displays

Stimuli were presented within a  $16.7 \times 16.7$  degrees of visual angle (deg) gray square. The square was surrounded by a 0.25 deg thick black and white border (Figure 1). The border encourages proper binocular fusion. Vergence was not monitored, however, and so vergence errors (described in the Introduction) will result in their eyes being only approximately converged at the correct angle. A  $10.7$  deg square region within the stimulus was populated with black and white dots. The dots were isotropic log-Gabors (peak spatial frequency 0.5 cycles/deg with a bandwidth of 2.4 octaves). The dots had a full width at half height of 7.5 arcmin. The dots were generated on a grid with a mean dot-to-dot spacing of 26 arcmin. To break up the structure of the grid, the dot positions were jittered in both the  $x$  and  $y$  dimensions with a random shift of up to  $\pm 11$  arcmin (drawn from a uniform distribution).

Disparity was introduced into the stimulus by shifting the horizontal position of the dots shown to each eye. The dots were shifted by an equal amount in opposite directions. The direction in which the dots were shifted determined whether the perceived

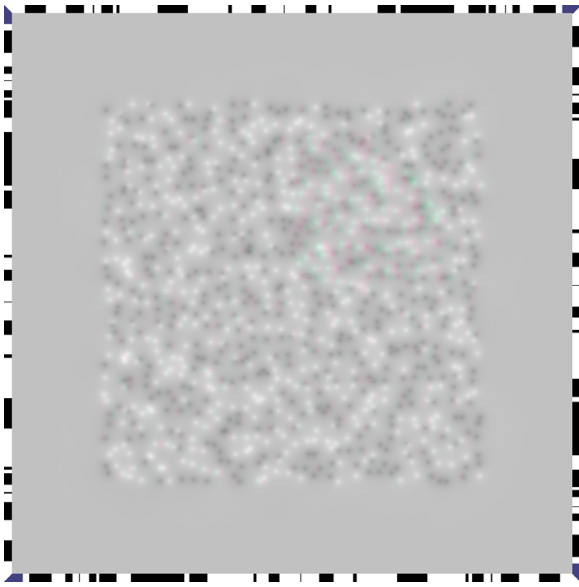


Figure 1. Example of the stimulus generated as a stereo anaglyph image that can be viewed with red-green glasses by the reader. At our viewing distance, this stimulus had a target disparity of 512 arcsec with 128 arcsec disparity noise standard deviation.

dot would have a crossed or an uncrossed disparity. Disparities not resolved in integer pixel shifts were achieved using subpixel interpolation. The disparity was used to introduce a target in one quadrant of the dot field. The target was a quarter-circle wedge with a radius of 4.5 deg. This wedge would appear either in front of (crossed) or behind (uncrossed) the dot field at the fixation plane. The inner corner of the wedge was positioned 8.5 arcmin from the center of the display. The wedge could appear in one of four places, the top-left, top-right, bottom-left, or bottom-right of the display. For large disparities, the horizontal shift of the dots in each eye created blank “voids” at the edges of the wedge stimulus. These voids could provide a monocular cue that would allow subjects to perform the task without perceiving depth. During stimulus generation, these regions were filled with dots at the fixation plane. Monocular control tests were conducted to ensure subjects needed to use the binocular disparity in the stimulus to perform the task.

For Experiment 2, stimulus disparity noise was introduced by randomly varying the disparity of each dot. The dot disparities were drawn from a normal distribution. The mean disparity was set to the target disparity within the target wedge. Elsewhere in the display, the mean disparity was zero. The standard deviation of the normal distribution determined the level of disparity noise. To reduce stimulus generation time, dots were pregenerated at a quantized list of disparities (log-spaced from 0.125 to 4096 arcsec in steps of 9%). When each dot was placed, its intended

disparity was drawn from a normal distribution. The closest disparity was then selected from the list of pregenerated dots and placed in the stimulus.

## Procedure

Thresholds for detecting the target were measured using a four-alternate forced-choice task. Subjects indicated which quadrant contained the wedge by pressing one of four keys. Subjects were given auditory feedback. Different pitches of beep were used to indicate correct or incorrect responses. The direction of disparity (crossed or uncrossed) was randomly selected on each trial. The magnitude of the disparity was determined by a pair of staircase routines for each direction (Baldwin, 2019). The staircases controlled the disparity on each trial, based on the subject’s performance. The staircases made the task more difficult when subjects responded correctly and easier when they responded incorrectly. The two randomly interleaved staircases for each disparity direction had different rules. One was a one-up-one-down staircase (converging at approximately 50% correct), and the other was a one-up-two-down staircase (converging at approximately 70% correct). Data were combined across the two staircases, and psychometric functions were fit to find thresholds. The initial disparity was set at 181 arcsec. Prior to the first reversal, the step size of the staircase was factors of two. After the first reversal, the step size reduced to a factor of  $\sqrt{2}$ . The experiment ended when all staircases reached either 50 total trials or eight reversals. The first three trials of every run had additional luminance contrast cues to indicate the location of the target. This familiarized subjects with the task. The subject had to respond correctly on the three training trials to continue the experiment. Data from these trials were not included in the analysis.

We conducted three experiments using variations of this basic paradigm. Experiment 1 tested 53 subjects, measuring thresholds for detecting crossed and uncrossed disparity. To make the task as simple as possible, the stimulus was presented for an unlimited duration. The subject could search the screen by making eye movements. No emphasis was placed on speed of response. After the subject pressed a key to respond, the stimulus disappeared, and feedback was given. The experiment then proceeded to the next trial. Each testing session lasted 5 to 6 minutes. Testing was repeated on a separate day to assess reliability of the results. Subjects with a large difference between their thresholds measured on the two days were tested a third time. Only the two most recent scores were used for the analysis, leading us to reject seven of 106 measurements made.

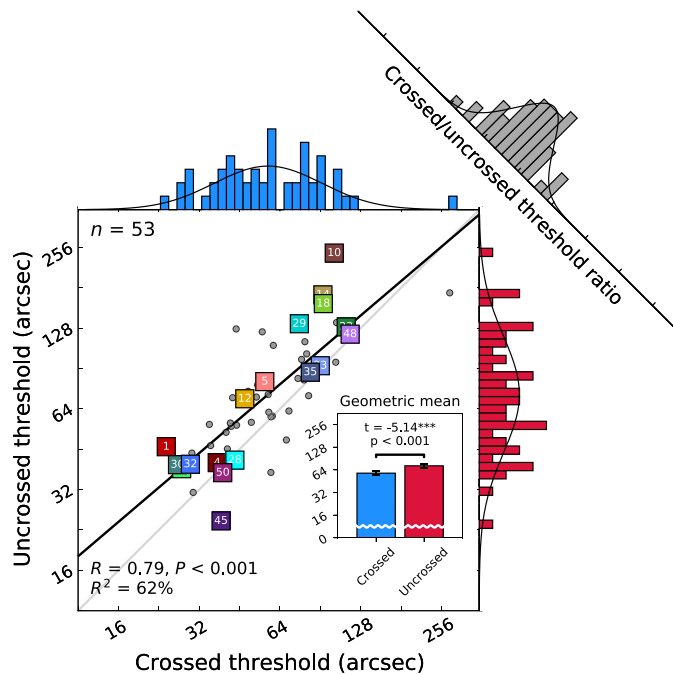


Figure 2. Correlation between thresholds for crossed and uncrossed disparity for 53 subjects. The 18 subjects who were selected for further testing are indicated with numbered square markers. The remaining subjects are indicated with gray circles. Blue (top) and red (right) histograms show the distribution of the crossed and uncrossed thresholds, respectively. The gray histogram (top-right) shows the distribution of the crossed/uncrossed threshold ratios. The inset bar plot shows the geometric mean crossed and uncrossed thresholds with their standard error. Data falling on the gray diagonal unity line have equal crossed and uncrossed thresholds. The black line is the best fit to the data:  $\log_2(y) = 0.85 \log_2(x) + 1.2$ .

In Experiment 2, we extensively tested 18 subjects (nine women) from the first experiment. The selected subjects gave a good coverage of the range of results from Experiment 1 (see numbered symbols in Figure 2). The stimulus duration was now fixed at 250 ms, to limit the duration subjects had to locate the target. The temporal envelope was a raised-cosine function. For the first 75 ms, the contrast ramped from zero to the full contrast (80%). It remained at this plateau for 250 ms, and then ramped back to zero over 75 ms. The subject could only respond once the stimulus had disappeared. We conducted a noise-masking experiment using these 250 ms stimuli. We tested subjects at seven external noise standard deviations. These were 0, 4, 32, 64, 128, 256, and 512 arcsec. The testing order of the noise levels was randomized. Subjects completed two repetitions for each condition. In cases in which the standard error of the threshold from the first two repetitions was unusually large, we collected additional data on that condition. This happened in nine of 252 measures made. The 512 arcsec

noise level for subject 48 was excluded from analysis as they could not be retested. In Experiment 3, we further tested 11 subjects in a control experiment in which the stimulus duration was reduced to 50 ms. The results from this experiment are reported in the Supplementary material. One subject (subject 30) only completed one repetition. Their results were not excluded.

## Analysis

Our performance measure was the threshold disparity for detecting the target in our four-alternative forced-choice task. Thresholds were estimated through logistic psychometric function fitting with Palamedes (Prins & Kingdom, 2018). We calculated the threshold corresponding to a correct response probability of 55.2%. This gives a signal-to-noise ratio ( $d'$ ) of unity at the decision variable. Crossed and uncrossed thresholds were obtained by combining data across both staircases for each disparity direction. We fit the psychometric function to the combined data for each direction. We also obtained an overall measure of stereo performance by fitting to the merged crossed and uncrossed data. We performed parametric bootstrapping for each psychometric function fit. We generated 500 bootstrap simulations to give us 500 bootstrapped threshold estimates. These were then used to calculate the standard error of the threshold.

In Experiments 2 and 3 (Supplementary material) we performed likelihood ratio tests in Palamedes (using the PAL\_PFLR\_ModelComparison function). These allowed us to determine whether crossed and uncrossed disparity data had the same threshold (Kingdom & Prins, 2016). This method uses Monte Carlo simulations (2000 samples) to calculate  $p$  values for the comparison. It therefore relies on stable psychometric function fits being made to the simulated data. In our usage, the fits to the simulations were not sufficiently stable to allow calculation of a  $p$  value in three out of 29 cases. For these we present approximate  $p$  values. We calculated these from the transformed likelihood ratios using the chi-squared survival function in the SciPy Python package (Virtanen, 2020).

## Equivalent noise analysis

We fit our data with the linear amplifier model

$$\text{threshold} = \frac{\sqrt{\sigma_{\text{external}}^2 + \sigma_{\text{eq.int.}}^2}}{\beta}, \quad (1)$$

where  $\sigma_{\text{external}}$  is the standard deviation of the disparity noise added to the stimulus,  $\sigma_{\text{eq.int.}}$  is the equivalent internal noise in the visual system of the subject, and

$\beta$  is a measure of how efficiently the subject is able to integrate the noisy information to perform the task. Noise-masking data are accounted for by varying the equivalent internal noise and efficiency parameters. These two parameters fit a threshold versus external noise curve to the data (plotted on log-log axes). Changes in the efficiency parameter translate the curve vertically. Changes in the equivalent internal noise move the transition point between the flat and sloped region horizontally.

Noise-masking functions were fit to the data for each subject and disparity direction. We performed the fitting in MATLAB (Mathworks, Natick, MA) using the `fminsearch` function. We fit the linear amplifier model by minimizing the root-mean-square error between the data and the model on a  $\log_2$ -threshold axis. We also obtained bootstrapped estimates of the standard error for the model parameters. We did this by fitting the linear amplifier model to the bootstrap threshold samples generated in the psychometric function fitting.

## Results

Crossed and uncrossed relative disparity thresholds measured in Experiment 1 are presented in Figure 2. The gray dots show thresholds from subjects who were not selected for further testing. The square symbols are used for subjects who were further tested in subsequent experiments. Thresholds ranged from 24 to 275 arcsec. Overall, data appear to lie along the diagonal. This indicates a rough agreement between subjects' crossed and uncrossed thresholds. To further explore this relationship, we performed a Pearson correlation analysis in SciPy (E. Jones, Oliphant, & Peterson, 2001). Thresholds for the two relative disparity directions were significantly correlated ( $R = 0.79$ ,  $p < 0.001$ ). Most of the points in Figure 2 lie *above* the unity line. This shows a tendency for uncrossed thresholds to be higher than those for crossed disparity. We tested for a mean difference using a  $t$ -test (SciPy) and found it to be significant:  $t(52)$ ,  $-5.14$ ,  $p < 0.001$ . The geometric mean was 73 arcsec for uncrossed thresholds and 58 arcsec for crossed thresholds. We performed a further analysis on the separate data from the two task repetitions for each subject. From this we found there to be a good test-retest reliability ( $R = 0.76$ ).

The noise-masking functions obtained in Experiment 2 are presented in Figure 3. The results from each subject are shown in separate panels. Thresholds are plotted separately for relative crossed and uncrossed disparity. A pair of thresholds are plotted at each external noise standard deviation. Performance for the two directions of disparity without noise (leftmost data point in each plot) were compared using a likelihood ratio test. The results of this test are presented in

Table S1 of the Supplementary material. Subjects 10, 12, 14, 28, 32, and 35 had significant threshold differences for detecting relative crossed versus uncrossed disparity (without external noise). As the external noise level increases, the classic noise masking behavior can be seen. With low levels of external noise, thresholds were constant. Beyond some critical value, however, thresholds increase linearly with external noise standard deviation. Thresholds were fit with the linear amplifier model. Most subjects who showed significant differences without external noise have fitted curves that are vertical translations of each other. This suggests that threshold differences are due to differences in the efficiency of processing for the two directions. There was also evidence, however, of differences in the transition point between the flat and sloped regions of the two curves. This indicates differences in equivalent internal noise.

We performed a series of analyses on the two parameters obtained from the linear amplifier model fits. We found no significant correlation between the equivalent internal noise and efficiency parameters across our 18 subjects. The  $R^2$  scores were 2% for the uncrossed disparity data ( $p = 0.622$ ), and 13% for the uncrossed disparity data ( $p = 0.150$ ). This suggests that the two parameters do not reflect a single underlying property that varies across our subjects. The noise-masking measurements allow us to ask whether either of the two model parameters are responsible for a larger part of the individual differences in disparity threshold. We analyzed the relationship between the relative disparity thresholds and the fitted model parameters (Figure 4). We took the thresholds from Experiment 1 and the linear amplifier model parameters from Experiment 2. This meant that the measurement errors affecting the  $x$ - and  $y$ -axes were independent. We used this analysis to investigate how the individual differences in stereo measured in Experiment 1 depend on the factors accounted for by the two model parameters. We performed a separate analysis for sensitivity to crossed (Figures 4A and B) and uncrossed (Figure 4C and D) relative disparity. Higher thresholds were associated with increased equivalent internal noise (Figures 4A and C) and reduced processing efficiency (Figures 4B and D). For each disparity direction we performed a multiple linear regression (using the ordinary least squares function from the statsmodels Python package; Seabold & Perktold, 2010). The regression predicted each subject's thresholds from Experiment 1 using the internal noise and efficiency parameters obtained from Experiment 2. The detailed results are presented in the Supplementary material. For both the crossed and uncrossed disparity results we found that subject-by-subject variation in both equivalent internal noise and processing efficiency contributed to the individual differences. The two parameters explained 73% of the variance for crossed

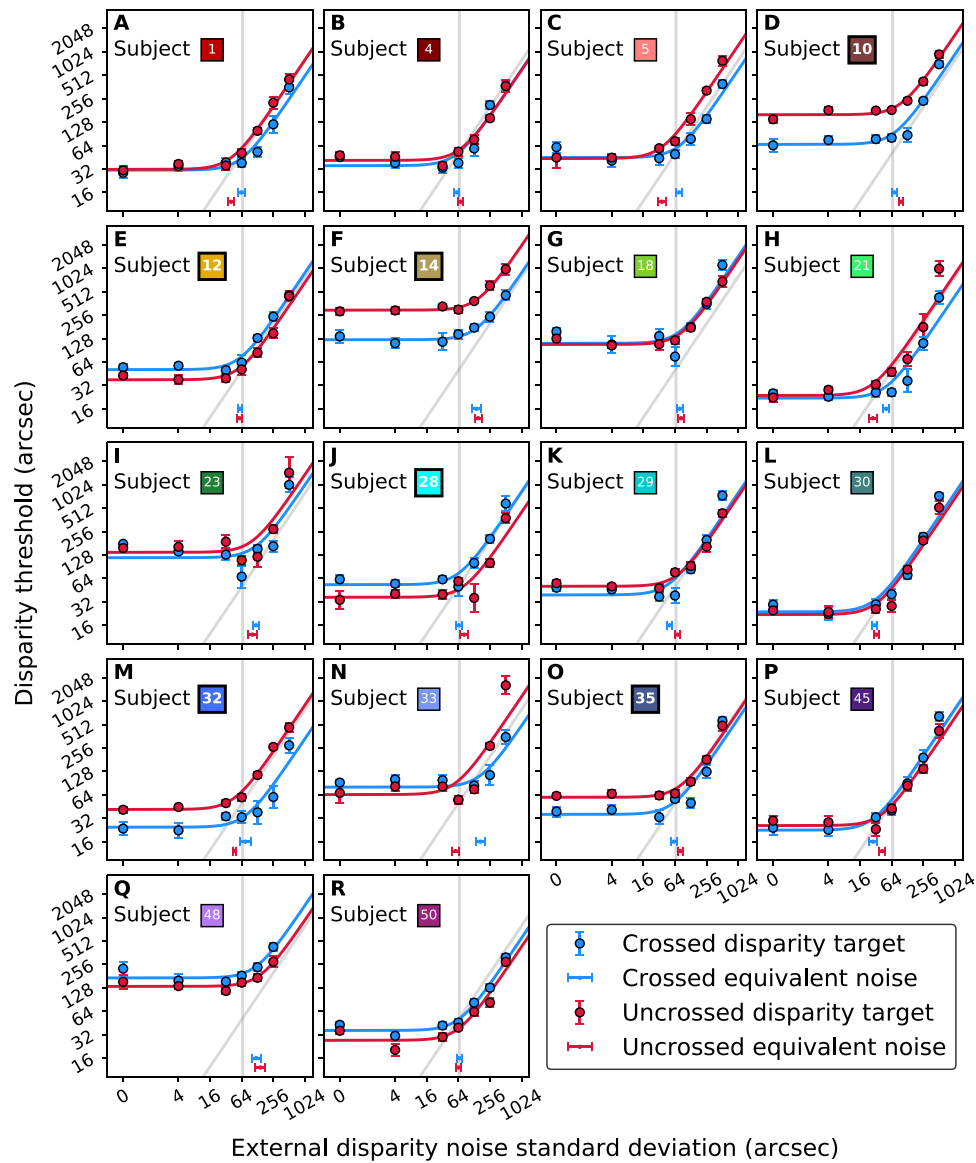


Figure 3. Noise-masking functions from Experiment 2 for 18 subjects in separate panels. Each panel includes the subject’s symbol from Figure 1. Thresholds ( $d' = 1$ ) for crossed and uncrossed disparities are plotted (with bootstrapped standard error) and fit with the linear amplifier model. The diagonal gray line shows the average efficiency across all subjects and conditions. The vertical gray line shows the average internal noise. The small symbols with horizontal error bars show the equivalent internal noise for crossed and uncrossed disparity for each subject.

disparity. For uncrossed disparity they explained 79% of the variance. For crossed disparity there was an even contribution from the two factors. Each parameter uniquely accounted for 30% to 34% of the variance. For uncrossed disparity the differences in efficiency were responsible for more of the threshold variation. Adding efficiency to the equivalent internal noise prediction provided a greater increase in  $R^2$  compared with adding equivalent internal noise to the efficiency prediction (32% vs. 20%).

Having looked at model parameters within a disparity direction, we can now compare the parameters

between crossed and uncrossed relative disparity. Equivalent internal noise and processing efficiency for each disparity direction were calculated from the data shown in Figure 3. The parameters for the two directions are plotted against each other in Figure 5. Data points on the gray diagonal unity line indicate equal values for crossed and uncrossed disparity. Symbols falling below the unity line have higher parameter values for crossed disparity. Those above the line indicate higher uncrossed disparity values. The equivalent internal noise parameters for the two directions were correlated, but the efficiencies were not.

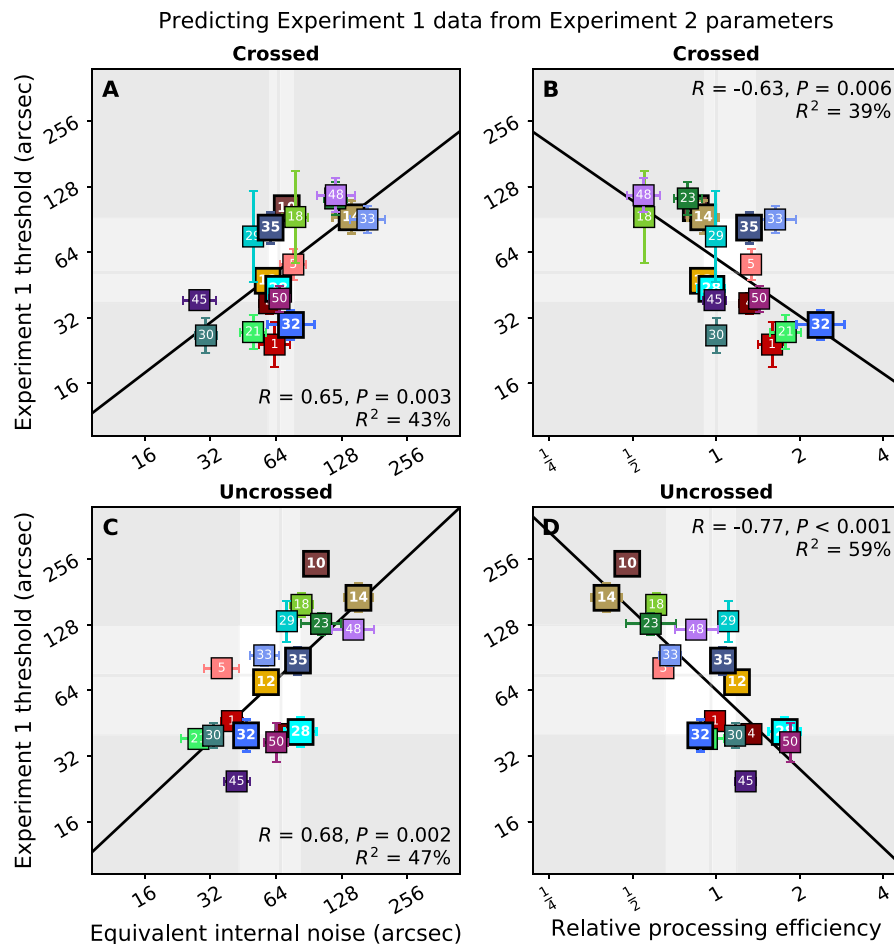


Figure 4. Correlations when using the parameters from Experiment 2 to predict the data from Experiment 1. The central gray cross splits the data about the median on each axis. Shaded regions at the beginning and end of each axis indicate the upper and lower quartiles. (A and C) show the correlation between equivalent internal noise (from Experiment 2) and threshold (from Experiment 1) for crossed and uncrossed disparity, respectively. (B and D) show the correlations between efficiency and threshold for the two disparity directions.

The symbol locations in Figure 5 can explain differences between the crossed and uncrossed disparity data in Figure 3. Most subjects with significant differences in thresholds without external noise had roughly equal equivalent internal noise for the two directions (Figure 5A). The only exception was subject 32. For all other subjects showing differences in sensitivity to crossed and uncrossed disparity, these were explained by efficiency differences (Figure 5B). For subject 32, it seems a combination of both factors is responsible for the difference in sensitivity.

The two parameters in the linear amplifier model have opposing effects on the noiseless threshold. A greater equivalent internal noise for one of the two relative disparity directions will result in an increase in threshold for that direction. This increase in noise could, however, be counteracted by a reciprocal increase in efficiency. This may result in no net change. Several subjects who did not have noiseless threshold

differences between crossed and uncrossed disparity are far from the unity line in Figures 5A and B. We hypothesize that these subjects may have balanced reciprocal differences in both equivalent internal noise and efficiency. We analyzed crossed:uncrossed equivalent internal noise and efficiency parameter ratios for each subject (Figure 6). If the two ratios were calculated based on the same data, we would expect some correlation between them. Any measurement error associated with the thresholds will affect the fitted parameters. Because the parameters interact, this error will have correlated effects on them both. To avoid this, we split the data into two subsets. We went through the list of noise mask levels and assigned each to either the odd set (0, 32, 128, and 512 arcsec) or the even set (4, 64, and 256 arcsec). We fit the linear amplifier model to each set. This gave us the equivalent internal noise and efficiency parameters for the two sets. We first compared the parameters obtained from fitting



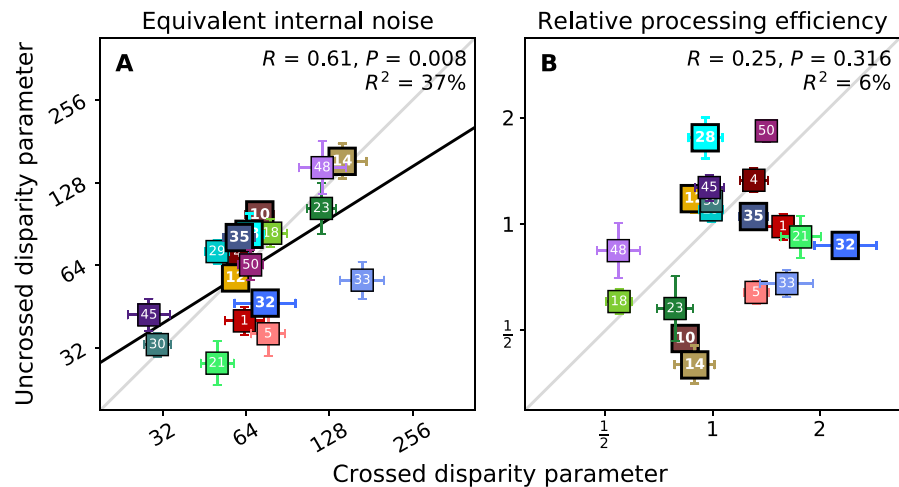


Figure 5. Correlation of linear amplifier model (LAM) parameters between the two disparity directions. (A) shows the comparison for equivalent internal noise. The diagonal gray unity line indicates equality between crossed and uncrossed parameters. Symbols are plotted with standard errors on each axis calculated from bootstrapping. The black line shows the best linear fit to the subject parameters:  $\log_2(y) = 0.63 \log_2(x) + 2.1$ . (B) shows the correlation between crossed and uncrossed efficiency. Heavier markers indicate subjects with a significant difference in crossed and uncrossed thresholds in Experiment 2 (from Table S1).

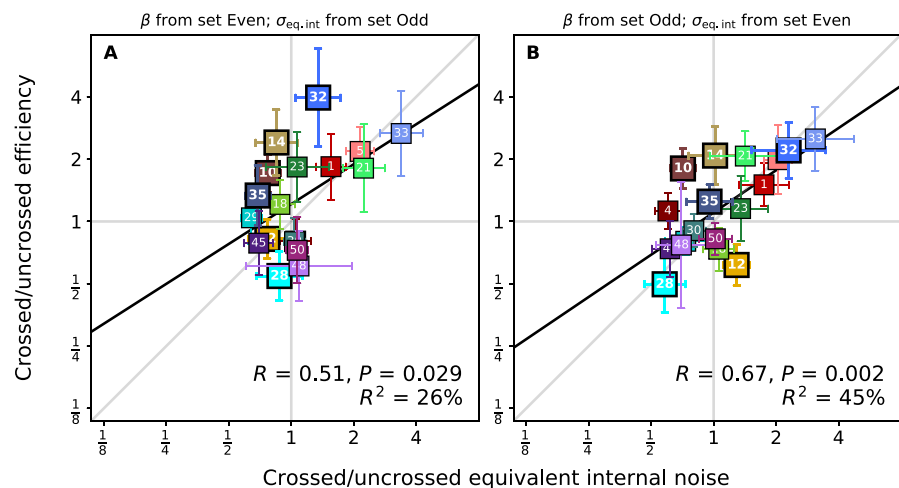


Figure 6. Relationship between crossed/uncrossed equivalent internal noise ratio for each subject and their crossed/uncrossed efficiency ratio. Subjects at (1, 1) have equal equivalent internal noise and efficiency for the two disparity directions. Subjects positioned along the diagonal gray unity line have shifts in both parameters with reciprocal effects, resulting in no net threshold change. To prevent spurious correlations from shared measurement error, the data were split into two subsets (odd and even) with one parameter obtained from the fit to the odd set compared with the other parameter obtained from the fit to the even set. The two plots (A and B) show this done in both possible ways (efficiency from set even/odd vs. equivalent internal noise from set odd/even).

to the data from the two sets. This was to verify that the parameters obtained from fitting to the odd set were similar to those obtained by fitting to the even set. There was a highly significant correlation for both equivalent internal noise ( $R^2 = 59\%$ ,  $p < 0.001$ ) and for processing efficiency ( $R^2 = 80\%$ ,  $p < 0.001$ ). This further serves as a validation of our modeling approach. Fits to data at the odd noise levels are highly predictive of the results at even noise levels. We then examined the

correlation between each parameter from the odd set with the other parameter from the even set. These ratios were significantly correlated. Individuals with increased equivalent internal noise for one disparity direction also tended to have higher processing efficiency for that direction. We can therefore split subjects into three groups: (i) those without significant differences between crossed and uncrossed disparity processing, (ii) those with balanced shifts in linear amplifier model

(LAM) parameters that yield no net change in noiseless thresholds, and (iii) those whose unbalanced shifts result in threshold differences between crossed and uncrossed disparity.

We performed additional control experiments to assess the possible role of stimulus duration, and refractive error on our results. These are reported in full in our Supplementary material. We found that reducing the stimulus duration from 250 ms to 50 ms (Experiment 3) did not change our findings. We also found a correlation between binocular visual acuity and stereoacuity, in agreement with previous reports (Bosten et al., 2015). Surprisingly, we found no significant correlation between equivalent internal noise and binocular visual acuity. Instead, there was a significant positive correlation between processing efficiency and binocular visual acuity. Subjects with better visual acuity were better at processing the noisy disparity information. To investigate this relationship, we retested three participants without their optical correction. This allowed us to see the effect of reduced visual acuity. Several studies have shown a reduction of stereoacuity on disturbance of visual acuity (Costa, Moreira, Hamer, & Ventura, 2010; Hess, Hong Liu, & Wang, 2002; Odell, Hatt, Leske, Adams, & Holmes, 2009). We found most of the effect of decreasing acuity was to increase equivalent internal noise. There was also a relatively modest reduction in processing efficiency.

## Discussion

Previous studies have shown substantial individual differences in the ability to perceive depth from disparity (Hess et al., 2015; Hess et al., 2016; Richards, 1970; Richards, 1971). It has been suggested that as many as one-third of the normal population may be stereoanomalous. This proportion has been shown to depend on the task used however (Dorman & van Ee, 2017; Hess et al., 2015, Hess et al., 2016; Landers & Cormack, 1997; Richards, 1970; Richards, 1971; van Ee & Richards, 2002; Wilcox, Hartle, Solski, Mackenzie, & Giaschi, 2017). We did not find such large individual differences with the task used in this study. In search of deficits specific to one disparity direction (Richards, 1971), we measured thresholds *separately* for targets in crossed and uncrossed disparity relative to the background. We found only modest differences between the two directions within a single subject. No subjects were “blind” to either direction of disparity (Richards, 1970; Richards, 1971). The three subjects with the greatest difference between the two directions were all better at detecting crossed disparity. Their uncrossed thresholds were between two and three times those measured for crossed disparity. They contributed to a significant asymmetry in sensitivity to the two

directions. This difference was consistent with previous literature (e.g., Schor & Wood, 1983).

The most dramatic examples of stereoanomaly found in previous studies have been from those that looked at coarse disparity (R. Jones, 1977; Richards, 1970; Richards & Regan, 1973). In this study, we are instead investigating fine disparity. Deficits for crossed or uncrossed absolute disparity may be disguised by subjects making vergence eye movements (Dorman & van Ee, 2017). This would allow a task testing a “crossed” stimulus to be transformed into a test of uncrossed stereoacuity. This could occur either as part of the subjects’ strategy, or as a consequence of the vergence error in the subjects’ eye movements (Ukwade et al., 2003). In Experiment 1, subjects were given an unlimited stimulus duration. These measurements would be especially vulnerable to a strategy based on deliberate eye movements. For this reason, our targets are defined as crossed or uncrossed relative to a reference plane. In Experiments 2 and 3 we presented stimuli with limited durations. In Experiment 2, stimuli were presented with a 250 ms duration. This did not reveal a difference greater than 3x between thresholds (without disparity noise) for crossed and uncrossed disparity. We further reduced the stimulus duration to 50 ms in Experiment 3 (reported in Supplementary material). In a sample of 11 subjects, the largest difference between crossed and uncrossed threshold was a factor of 1.8x. It still remains possible, however, that the smaller differences between crossed and uncrossed disparity in our fine disparity task are due to errors in vergence eye movements. These would randomly shift the disparity of the entire stimulus arrangement. This may blur the effects of any specific deficit to detecting crossed or uncrossed absolute disparity. For the purpose of this study we restrict our conclusions to the discussion of relative disparity.

Another difference in our task is that we required subjects only to detect the presence of disparity. They did not have to discriminate between disparity directions. Previous studies finding larger incidences of stereoanomaly (Hess et al., 2015; Hess et al., 2016; Landers & Cormack, 1997; Richards, 1970; Richards, 1971; van Ee & Richards, 2002; Wilcox et al., 2017) have asked subjects to judge the direction of disparity. Studies in which subjects simply detect the presence of disparity typically find narrower ranges of ability (e.g., Bohr & Read, 2013). Narrower ranges have also been reported from some studies in which subjects judged the direction of disparity (Coutant & Westheimer, 1993; Zaroff et al., 2003). In this study we found a range from 24 to 275 arcsec. It could be that our testing would have found larger individual differences if we had subjects discriminate the direction of disparity. One explanation would be if the identification of the direction of disparity happens separately to detection (Landers & Cormack, 1997). This ability might be selectively

impaired in stereoanomaly. This deficit may be hidden in normal visual function. The presence of other cues to depth (e.g. motion parallax) could disambiguate the direction of disparity.

Although we found a smaller degree of variability than some previous studies, our subjects still showed an elevenfold range in stereoacuity. We also found significant within-subject differences in sensitivity to crossed and uncrossed disparity. In the most extreme cases these approached a factor of three. Different explanations have been put forth to explain the asymmetries in disparity sensitivity. These include disturbed development of one system (Dorman & van Ee, 2017; Richards, 1971) and vergence anomalies (R. Jones, 1977). Specific to uncrossed stimuli, a lack of well-defined boundaries for uncrossed random dot stimuli has been suggested to hinder performance (Becker, Bowd, Shorter, King, & Patterson, 1999).

We investigated these stereoacuity differences further using the noise masking paradigm (Experiment 2). Fitting the data using the linear amplifier model allowed us to break down the individual differences into equivalent internal noise and processing efficiency factors. Both model parameter factors from Experiment 2 were correlated with Experiment 1 thresholds. We concluded that subject-by-subject variations in both processing efficiency and equivalent internal noise contributed to individual differences in stereoacuity. We did find evidence for a slightly larger role of efficiency in accounting for the individual differences with uncrossed stimuli. Several subjects who lack threshold differences between crossed and uncrossed disparities do show an imbalance in their noise-masking parameters between the two directions. This could be explained by the visual system compensating for a noisier input by adjusting the efficiency of processing. It should be noted that this analysis was conducted within the framework of the linear amplifier model. We chose this approach as it provides a simple method for analyzing our noise-masking data. Our data appear to agree with the behavior expected according to that model. It is almost certain that a full account of stereo sensitivity will require a model that is significantly more complex than the one we apply here. Any such model will, however, have to replicate the behavior of the linear amplifier model for the purpose of accounting for these results.

## Conclusions

The goal of this study was to understand the underlying cause of the range of stereoacuity ability in the normal population. We used a simple relative disparity detection task. Subjects located stimuli that were rendered in a different disparity to a background. Our stimulus design was therefore similar to that

used in clinical testing (Birch et al., 2008), although the details of the task differed. We found an 11-fold range in stereoscopic sensitivity. We also found a bias in processing the two disparity directions, with an advantage for crossed disparity. Equivalent internal noise levels (reflecting the signal to noise ratio of the disparity signals) and processing efficiencies (reflecting how well that noisy information is processed) were calculated. We found that both factors contribute to the spread of stereo abilities in our sample. This suggests that individual differences in central processing efficiency and equivalent internal noise are both important factors contributing to the range in stereoacuity.

*Keywords:* psychophysics, binocular disparity, noise masking, depth perception, stereoacuity

## Acknowledgments

Supported by research grants from the Natural Sciences and Engineering Research Council of Canada (awarded to RFH #228103), and the Vision Health Research Network of the *Fonds de Recherche du Québec-Santé* (awarded to SAC).

Data will be made available on figshare.com, with reserved DOI: 10.6084/m9.figshare.7614140.

Commercial relationships: Alex S. Baldwin and Robert F. Hess are inventors on a PCT patent application “System and Method for Digital Measurement of Stereo Vision” (CA2020050051, filed January 17th 2020), which has been commercially licensed by McGill University.

Corresponding author: Alex S. Baldwin.  
Email: alexsbaldwin@googlemail.com.  
Address: McGill Vision Research, McGill University, Montreal, Quebec, Canada.

## References

- Bach, M. (1996). The Freiburg visual acuity test: Automatic measurement of visual acuity. *Optometry and Vision Science*, 73(1), 49–53, <https://michaelbach.de/fract/>.
- Baldwin, A. S., Fu, M., Farivar, R., & Hess, R. F. (2017). The equivalent internal orientation and position noise for contour integration. *Scientific Reports*, 7, 13048, <https://doi.org/10.1038/s41598-017-13244-z>.
- Baldwin, A. S. (2019). *alexbaldwin/MatlabStaircase: v0.9.0*. Retrieved from: <https://doi.org/10.5281/zenodo.3266142>.

- Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *Journal of Physiology*, *193*, 327–342.
- Becker, S., Bowd, C., Shorter, S., King, K., & Patterson, R. (1999). Occlusion contributes to temporal processing differences between crossed and uncrossed stereopsis in random-dot displays. *Vision Research*, *39*, 331–339, [https://doi.org/10.1016/S0042-6989\(98\)00110-2](https://doi.org/10.1016/S0042-6989(98)00110-2).
- Birch, E. E., Gwiazda, J., & Held, R. (1982). Stereoacuity development for crossed and uncrossed disparities in human infants. *Vision Research*, *22*(5), 507–513, [https://doi.org/10.1016/0042-6989\(82\)90108-0](https://doi.org/10.1016/0042-6989(82)90108-0).
- Birch, E., Williams, C., Drover, J., Fu, V., Cheng, C., & Northstone, K., ... Adams, R. (2008). Randot Preschool Stereoacuity Test: Normative data and validity. *Journal of AAPOS*, *12*(1), 23–26, <https://doi.org/10.1016/j.jaapos.2007.06.003>.
- Bohr, I., & Read, J. C. A. (2013). Stereoacuity with Frisby and revised FD2 stereo tests. *PLoS One*, *8*(12), e82999, <https://doi.org/10.1371/journal.pone.0082999>.
- Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Bargary, G., Hogg, R. E., & Mollon, J. D. (2015). A population study of binocular function. *Vision Research*, *110*(Part A), 34–50, <https://doi.org/10.1016/j.visres.2015.02.017>.
- Cohn, T. E. (1976). Detectability of a luminance increment: Effect of superimposed random luminance fluctuation. *Journal of the Optical Society of America A*, *66*(12), 1426–1428.
- Cornell, E. D., Macdougall, H. G., Predebon, J., & Curthoys, I. S. (2003). Errors of binocular fixation are common in normal subjects during natural conditions. *Optometry and Vision Science*, *80*(11), 764–771, <https://doi.org/10.1097/00006324-200311000-00014>.
- Costa, M. F., Moreira, S. M. C. F., Hamer, R. D., & Ventura, D. F. (2010). Effects of age and optical blur on real depth stereoacuity. *Ophthalmic and Physiological Optics*, *30*, 660–666, <https://doi.org/10.1111/j.1475-1313.2010.00750.x>.
- Coutant, B. E., & Westheimer, G. (1993). Population distribution of stereoscopic ability. *Ophthalmic and Physiological Optics*, *13*, 3–7, <https://doi.org/10.1111/j.1475-1313.1993.tb00419.x>.
- Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America A*, *18*(5), 1016–1026, <https://doi.org/10.1364/JOSAA.18.001016>.
- DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1991). Depth is encoded in the visual cortex by a specialised receptive field structure. *Nature*, *352*, 156–159.
- Dorman, R., & van Ee, R. (2017). 50 Years of stereoblindness: Reconciliation of a continuum of disparity detectors with blindness for disparity in near or far depth. *I-Perception*, *8*(6), 1–13, <https://doi.org/10.1177/2041669517738542>.
- Fredenborg, P., & Harwerth, R. S. (2001). The relative sensitivities of sensory and motor fusion to small binocular disparities. *Vision Research*, *41*, 1969–1979, [https://doi.org/10.1016/S0042-6989\(01\)00081-5](https://doi.org/10.1016/S0042-6989(01)00081-5).
- Hess, R. F., Ding, R., Clavagnier, S., Liu, C., Guo, C., & Viner, C., ... Zhou, J. (2016). A robust and reliable test to measure stereopsis in the clinic. *Investigative Ophthalmology and Visual Science*, *57*, 798–804, <https://doi.org/10.1167/iovs.15-18690>.
- Hess, R. F., Hong Liu, C., & Wang, Y. Z. (2002). Luminance spatial scale and local stereo-sensitivity. *Vision Research*, *42*, 331–342, [https://doi.org/10.1016/S0042-6989\(01\)00285-1](https://doi.org/10.1016/S0042-6989(01)00285-1).
- Hess, R. F., Mansouri, B., Dakin, S. C., & Allen, H. A. (2006). Integration of local motion is normal in amblyopia. *Journal of the Optical Society of America A*, *23*(5), 986–992.
- Hess, R. F., To, L., Zhou, J., Wang, G., Cooperstock, J. R., & Long, T., ... Cooperstock, J. R. (2015). Stereo vision: The haves and have-nots. *I-Perception*, *6*(3), 1–5, <https://doi.org/10.1177/2041669515593028>.
- Jones, E., Oliphant, T., & P. (2001). *SciPy: Open source scientific tools for Python*. Retrieved from [https://scipy.github.io/old-wiki/pages/Citing\\_SciPy.html](https://scipy.github.io/old-wiki/pages/Citing_SciPy.html).
- Jones, R. (1977). Anomalies of disparity detection in the human visual system. *The Journal of Physiology*, *264*(3), 621–640, <https://doi.org/10.1113/jphysiol.1977.sp011686>.
- Julesz, B. (1960). Binocular depth perception of computer-generated patterns. *The Bell System Technical Journal*, *39*(5), 1125–1162.
- Kingdom, F. A. A., & Prins, N. (2016). Model comparisons. In *Psychophysics: A practical introduction* (2nd ed.). (pp. 247–307). London: Academic Press.
- Krauskopf, J., Cornsweet, T. N., & Riggs, L. A. (1960). Analysis of eye movements during monocular and binocular fixation. *Journal of the Optical Society of America*, *50*(6), 572–578, <https://doi.org/10.1364/JOSA.50.000572>.
- Landers, D. D., & Cormack, L. K. (1997). Asymmetries and errors in perception of depth from disparity suggest a multicomponent model of disparity processing. *Perception and Psychophysics*, *59*(2), 219–231, <https://doi.org/10.3758/BF03211890>.

- Manning, M. L., Finlay, D. C., Neill, R. A., & Frost, B. G. (1987). Detection threshold differences to crossed and uncrossed disparities. *Vision Research*, 27(9), 1683–1686, [https://doi.org/10.1016/0042-6989\(87\)90174-X](https://doi.org/10.1016/0042-6989(87)90174-X).
- McKee, S. P. (1983). The spatial requirements for fine stereoacuity. *Vision Research*, 23(2), 191–198, [https://doi.org/10.1016/0042-6989\(83\)90142-6](https://doi.org/10.1016/0042-6989(83)90142-6).
- McKee, S. P., Welch, L., Taylor, D. G., & Bowne, S. F. (1990). Finding the common bond: Stereoacuity and the other hyperacuties. *Vision Research*, 30(6), 879–891, [https://doi.org/10.1016/0042-6989\(90\)90056-Q](https://doi.org/10.1016/0042-6989(90)90056-Q).
- Newhouse, M., & Uttal, W. R. (1982). Distribution of stereoanomalies in the general population. *Bulletin of the Psychonomic Society*, 20(1), 48–50, <https://doi.org/10.3758/BF03334799>.
- Odell, N. V., Hatt, S. R., Leske, D. A., Adams, W. E., & Holmes, J. M. (2009). The effect of induced monocular blur on measures of stereoacuity. *Journal of AAPOS*, 13, 136–141, <https://doi.org/10.1016/j.jaapos.2008.09.005>.
- Patterson, R., Cayko, R., Short, G. L., Flanagan, R., Moe, L., & Taylor, E., ... Day, P. (1995). Temporal integration differences between crossed and uncrossed stereoscopic mechanisms. *Perception & Psychophysics*, 57(6), 891–897, <https://doi.org/10.3758/BF03206803>.
- Patterson, R., & Fox, R. (1984). The effect of testing method on stereoanomaly. *Vision Research*, 24(5), 403–408, [https://doi.org/10.1016/0042-6989\(84\)90038-5](https://doi.org/10.1016/0042-6989(84)90038-5).
- Pelli, D. G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America A*, 16(3), 647–653, <https://doi.org/10.1364/josaa.16.000647>.
- Pelli, D. G. (1981). *Effects of visual noise*. Cambridge: Cambridge University.
- Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, 40(6), 1392–1405, <https://doi.org/10.1152/jn.1977.40.6.1392>.
- Prins, N., & Kingdom, F. A. A. (2018). Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox. *Frontiers in Psychology*, 9: 1250, doi:10.3389/fpsyg.2018.01250. <http://www.palamedestoolbox.org/>.
- Richards, W. (1970). Stereopsis and stereoblindness. *Experimental Brain Research*, 10(4), 380–388, <https://doi.org/10.1007/BF02324765>.
- Richards, W. (1971). Anomalous stereoscopic depth perception. *Journal of the Optical Society of America*, 61(3), 410–414, <https://doi.org/10.1364/JOSA.61.000410>.
- Richards, W., & Regan, D. (1973). A stereo field map with implications for disparity processing. *Investigative Ophthalmology*, 12(12), 904–909.
- Schor, C. M. (2013). Stereopsis. In Leo M. Chalupa, & John S. Werner (Eds.), *The new visual neurosciences* (pp. 809–823). Cambridge, MA: MIT Press.
- Schor, C. M., & Wood, I. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research*, 23(12), 1649–1654, [https://doi.org/10.1016/0042-6989\(83\)90179-7](https://doi.org/10.1016/0042-6989(83)90179-7).
- Seabold, S., & Perktold, J. (2010). Statsmodels: econometric and statistical modeling with Python. In *9th Python in Science Conference*, <https://doi.org/10.1007/s00367-011-0258-7>.
- Serrano-Pedraza, I., Vancleef, K., & Read, J. C. A. (2016). Avoiding monocular artifacts in clinical stereotests presented on column-interleaved digital stereoscopic displays. *Journal of Vision*, 16(14):13, 1–14 <https://doi.org/10.1167/16.14.13>.
- Steinman, R. M., Cushman, W. B., & Martins, A. J. (1982). The precision of gaze. *Human Neurobiology*, 1, 97–109.
- Stevenson, S. B., Cormack, L. K., & Schor, C. M. (1989). Hyperacuity, superresolution and gap resolution in human stereopsis. *Vision Research*, 29, 1597–1605, [https://doi.org/10.1016/0042-6989\(89\)90141-7](https://doi.org/10.1016/0042-6989(89)90141-7).
- Tittes, J., Baldwin, A. S., Hess, R. F., Cirina, L., Wenner, Y., & Kuhli-Hattenbach, C., ... Fronius, M. (2019). Assessment of stereovision with digital testing in adults and children with normal and impaired binocularity. *Vision Research*, 164, 69–82, <https://doi.org/10.1016/j.visres.2019.07.006>.
- Ukwade, M. T., Bedell, H. E., & Harwerth, R. S. (2003). Stereopsis is perturbed by vergence error. *Vision Research*, 43(2), 181–193, [https://doi.org/10.1016/S0042-6989\(02\)00408-X](https://doi.org/10.1016/S0042-6989(02)00408-X).
- Virtanen, P., Gommers, R., & Oliphant, T. E., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17, 261–272, <https://doi.org/10.1038/s41592-019-0686-2>.
- van Ee, R., & Richards, W. (2002). A planar and a volumetric test for stereoanomaly. *Perception*, 31, 51–64 <https://doi.org/10.1068/p3303>.
- Wardle, S. G., Bex, P. J., Cass, J., & Alais, D. (2012). Stereoacuity in the periphery is limited by internal noise. *Journal of Vision*, 12(6), 12, <https://doi.org/10.1167/12.6.12>.
- Webber, A. L., Schmid, K. L., Dowdall, R. J., Finlay, M. E., Larner, R. L., & Lee, E., ... Hess, R. F. (2018). Suppression rather than visual acuity loss limits

- threshold stereoacuity. *Investigative Ophthalmology & Visual Science*, *59*, 4115.
- Wheatstone, C. (1838). Contributions to the Physiology of Vision 1. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, *8*, 371–394, <https://doi.org/10.1017/cbo9781139057950.026>.
- Wilcox, L. M., Hartle, B., Solski, A., Mackenzie, K. J., & Giaschi, D. (2017). Disparity configuration influences depth discrimination in naïve adults, but not in children. *Vision Research*, *131*, 106–119, <https://doi.org/10.1016/j.visres.2016.12.010>.
- Zaroff, C. M., Knutelska, M., & Frumkes, T. E. (2003). Variation in stereoacuity: Normative description, fixation disparity, and the roles of aging and gender. *Investigative Ophthalmology & Visual Science*, *44*(2), 891–900, <https://doi.org/10.1167/iops.02-0361>.