

# SCIENTIFIC REPORTS



OPEN

## First- and second-order contributions to depth perception in anti-correlated random dot stereograms

Jordi M. Asher  & Paul B. Hibbard

The binocular energy model of neural responses predicts that depth from binocular disparity might be perceived in the reversed direction when the contrast of dots presented to one eye is reversed. While reversed-depth has been found using anti-correlated random-dot stereograms (ACRDS) the findings are inconsistent across studies. The mixed findings may be accounted for by the presence of a gap between the target and surround, or as a result of overlap of dots around the vertical edges of the stimuli. To test this, we assessed whether (1) the gap size (0, 19.2 or 38.4 arc min) (2) the correlation of dots or (3) the border orientation (circular target, or horizontal or vertical edge) affected the perception of depth. Reversed-depth from ACRDS (circular no-gap condition) was seen by a minority of participants, but this effect reduced as the gap size increased. Depth was mostly perceived in the correct direction for ACRDS edge stimuli, with the effect increasing with the gap size. The inconsistency across conditions can be accounted for by the relative reliability of first- and second-order depth detection mechanisms, and the coarse spatial resolution of the latter.

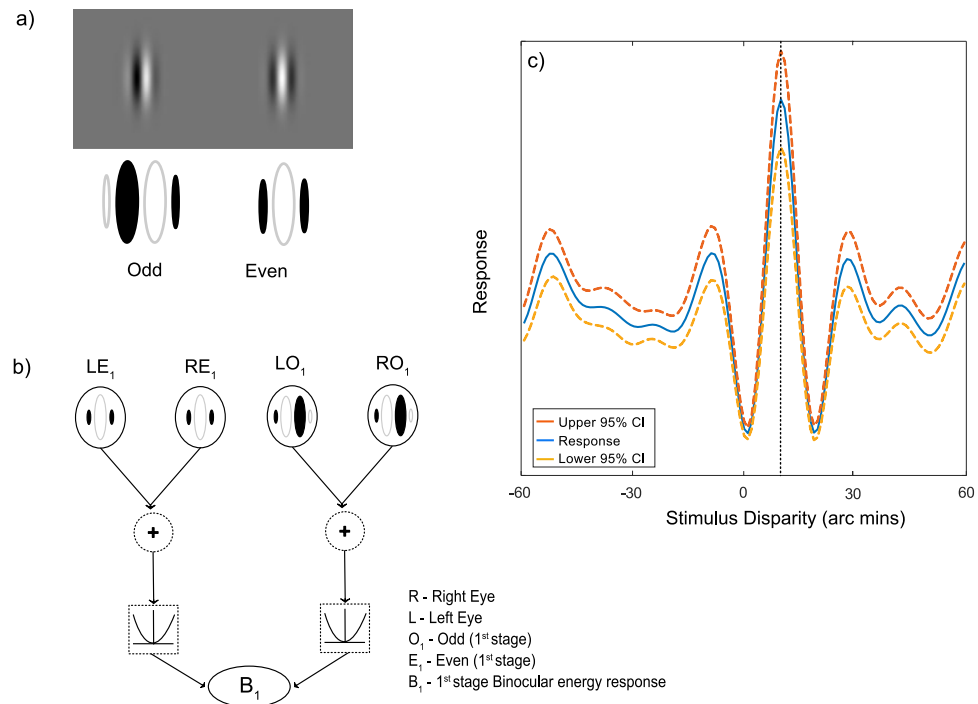
Binocular depth perception depends on our ability to determine the difference in position of corresponding points between the two eyes' images. This correspondence problem can be solved by using a matching measure to determine the image regions in each eye which are most similar in terms of the variation in local luminance intensity. This similarity matching can for example be based on the correlation of local intensity values<sup>1-4</sup>. At the correct disparity offset between the left and right eyes, each point is expected to have similar luminance values in each eye, leading to a high interocular correlation. At incorrect disparities, non-corresponding points will be compared, which are likely to have different values of luminance, leading to low correlations.

In addition to the standard correlation-based measures described above, other matching-based metrics have been proposed that depend on detecting similarities between two image samples, based on the presence of individual matching features, rather than the overall correlation within the region<sup>5,6</sup>. These metrics are not necessarily mutually exclusive, and it has been suggested that both metrics are used independently and simultaneously<sup>5,6</sup>. Under the matching metric proposed, all evidence in favour of a match, when points have the same luminance polarity, contributes positively to the matching metric. However, in contrast to a standard correlation, points that have opposite contrast polarities, and provide evidence against a match, are ignored.

These similarity calculations can be related to the responses of binocular neurons in the visual cortex<sup>3,4,7,8</sup>. Neurons in V1 have a localised receptive field, consisting of both excitatory and inhibitory regions, and tend also to be tuned to orientation and spatial frequency<sup>3,9-12</sup>. Binocular neurons have a receptive field in both eyes, and their responses are thus sensitive to binocular disparities. Figure 1(a) shows an idealised binocular energy neuron<sup>3,9,10</sup>, consisting of a quadrature pair of receptive fields for each eye. The responses of the first-stage filters are the square of the sum across the two eyes' receptive fields (Fig. 1(b)), which in turn are summed to compute the binocular energy response.

These components of the energy model have been used to characterise the responses of binocular simple and complex cells, respectively, although this simple hierarchy is best viewed as an idealisation of how these computations are performed<sup>13</sup>. The response of a binocular energy neuron depends on the disparity in the images, forming a characteristic disparity tuning function (Fig. 1(c)). In this example, the receptive fields of the filters are shifted

University of Essex, Department of Psychology, Wivenhoe Park, CO4 3SQ, United Kingdom. Jordi M. Asher and Paul B. Hibbard contributed equally. Correspondence and requests for materials should be addressed to J.M.A. (email: [jashera@essex.ac.uk](mailto:jashera@essex.ac.uk))



**Figure 1.** (a) Odd and even symmetric monocular receptive fields (b) These receptive fields are combined to compute the binocular energy response (c) The responses of binocular energy neurons show a characteristic tuning function. Here the response, calculated as the mean over 1000 Gaussian white noise samples, shows a peak at the disparity (10 arc min) corresponding to the difference in position of the monocular receptive fields. Dotted lines show upper and lower 95% confidence intervals.

horizontally in the right eye compared to those in the left eye, meaning that this neuron responds most strongly to stimuli with the same disparity. This peak response at this optimum disparity is accompanied by disparities at which the response is reduced compared to baseline.

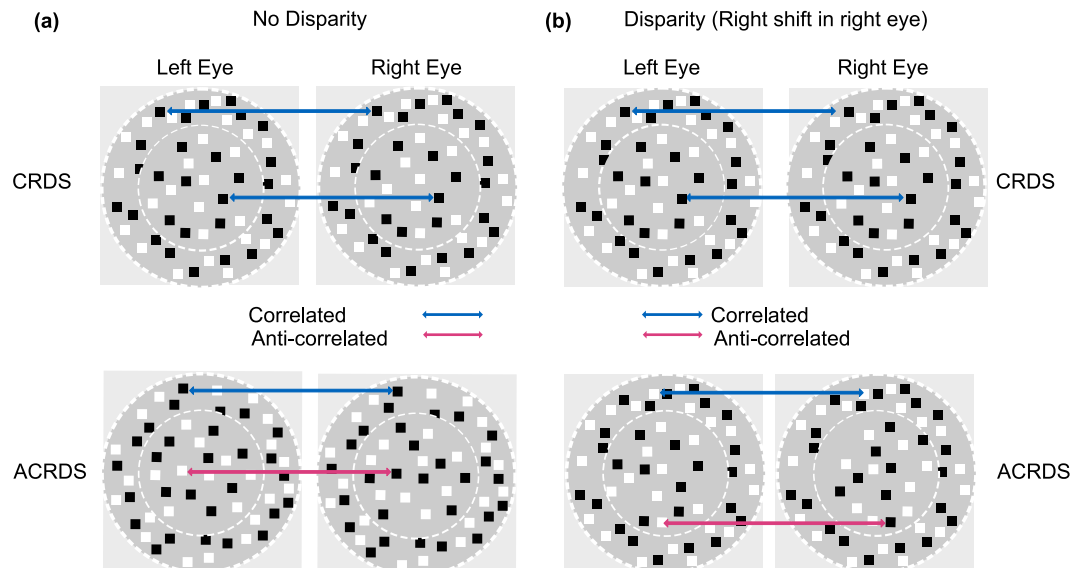
The binocular energy response is related to the point-wise correlation between the filtered left and right images<sup>3</sup>. By summing across frequency, orientation and position<sup>3,14,15</sup>, the correlation within a spatial neighbourhood can be calculated<sup>3,4</sup>. Complex cells in V1, as characterised by the binocular energy model, are thus well suited to support the calculation of the cross-correlation and other matching calculations thought to contribute to the solution of the correspondence problem<sup>9</sup>.

The dependence of the energy response on binocular cross-correlation means that manipulation of this correlation has provided a useful way of understanding how the responses of populations of binocular neurons contribute to the perception of depth<sup>2</sup>. This is often investigated psychophysically using random-dot stereograms (RDS) which, by projecting an image that has been shifted slightly between the left and right eye, creates the perception of depth. Correlated random-dot stereograms (CRDS) present dots of the same luminance to each eye (Fig. 2(a)). However, an interesting application of RDS is the use of anti-correlated random-dot stereogram stimuli (ACRDS), where one eye's view is replaced with its photographic negative (Fig. 2(b))<sup>5,12,16–23</sup>. This means that the high positive correlations expected at the correct disparity become negative, and the disparity tuning function is inverted. Neurons in V1 show this inversion effect, but also a reduction in magnitude of their response that is not predicted by the energy model<sup>9–11</sup>. This reduction in response has been modelled using the introduction of a threshold non-linearity<sup>24–26</sup>, or a squaring of the energy response<sup>27</sup>. These expansive nonlinearities, by enhancing the difference between the amplitudes of the positive and negative peaks in the disparity tuning function, can be used to implement the cross-matching mechanism proposed by Doi and Fujita<sup>5,6</sup>.

In higher visual areas in the ventral stream, the responses of neurons tend not to be modulated by the disparity in ACRDS<sup>28,29</sup>. In contrast, neurons in dorsal stream areas show disparity tuning similar to that found in V1, but reduced in magnitude<sup>30</sup>.

In some psychophysical studies, the direction of depth perceived in ACRDS has been found to be reversed in comparison with equivalent CRDS<sup>5,17,21,23</sup>. Thus, we use the terms 'reversed-depth' to describe depth that is perceived in the opposite direction to the correct direction, and 'forward-depth' depth that is perceived in the correct direction. One possible explanation of this percept is that it reflects the peaks in the inverted disparity tuning function, although whether these would signal a reversed- or forward-depth direction depends on the relationship between the spatial frequency and disparity tuning of the neuron, and the stimulus disparity (Fig. 3).

Alternatively, it has been suggested that the estimation of depth might reflect opponent processing, in which the difference between the responses of neurons tuned to equal but opposite disparities are calculated<sup>6,27</sup>. In this case, the negative correlations that exist in ACRDS would directly contribute to the reduction in the response to the correct disparity, thereby biasing the perception of depth towards the opposite direction.



**Figure 2.** (a) Correlated and Anti-correlated RDSs, left image is presented to the left eye and the right to the right eye. Both CRDS and ACRDS have a correlated reference (surround annulus). Correlated RDS (top) have correlated dots in the target (centre circle). In contrast, the central targets in anti-correlated RDS (bottom) have reversed luminance for the corresponding eyes. (b) Dots in the target (centre circle) include a rightwards shift for the right eye's image. The surround for all cases remains correlated with zero disparity.

Other studies have found no evidence for the perception of depth in ACRDS<sup>18,19,21</sup>. This result is consistent with the fact that there is no coherent disparity, across different scales of analysis, in ACRDS, and also with the lack of disparity-selective responses in higher cortical areas.

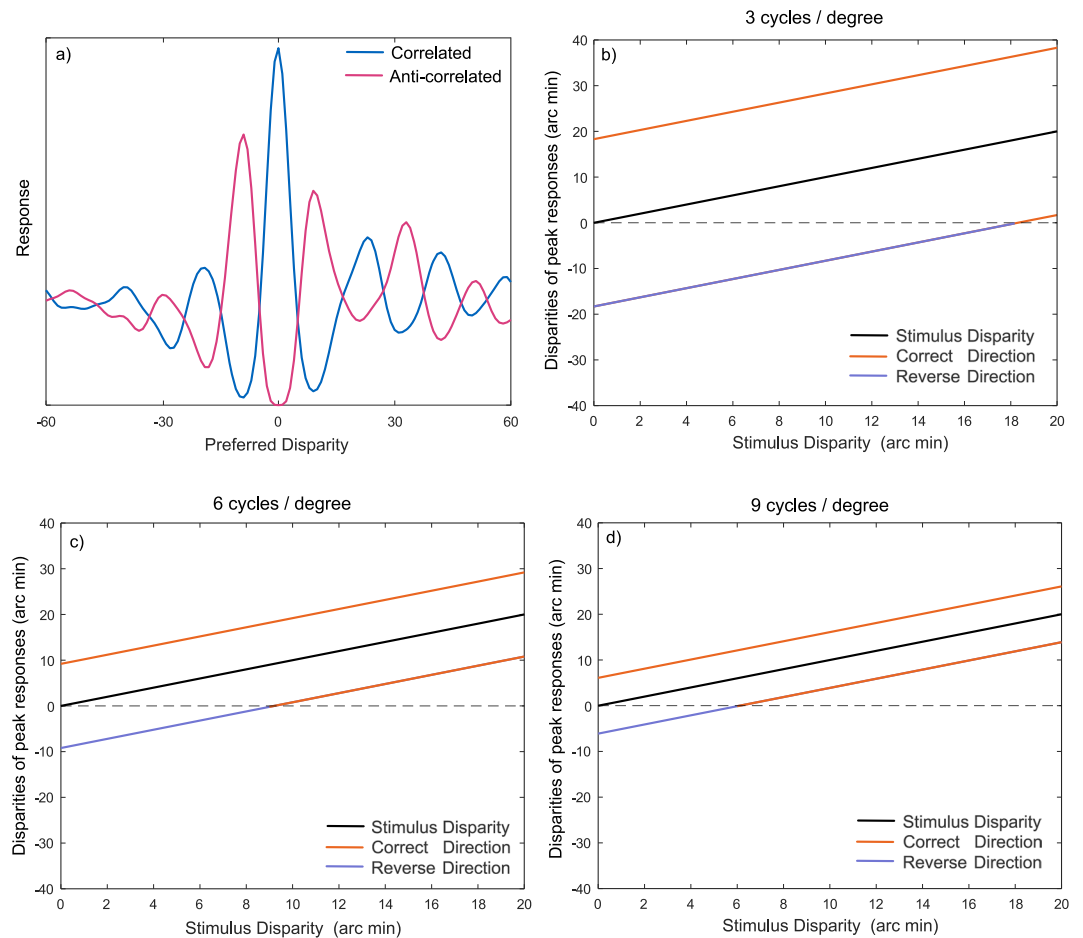
The effect of decorrelation has been further assessed by Doi *et al.*<sup>5</sup> who created stimuli containing an equal mixture of correlated and anti-correlated dot pairs. This results in an overall correlation of zero, so that a correlation-based mechanism predicts that depth would not be seen in these half-matched stimuli. In fact, depth is perceived in the correct direction. This is consistent with the responses of the cross-matching mechanism, which responds to the correct-matched dot pairs, but not to the anti-correlated pairs. Henriksen *et al.*<sup>27</sup> showed that a similar prediction can be made by the squared energy response, which enhances the difference between the response to the paired and unpaired dots.

Predicting the perception of depth in ACRDS is complicated by the fact that there is no coherent disparity signal across different spatial scales. In CRDS, pooling of information across orientation, scale and position allows true peaks to emerge from amongst the many large responses that will occur at incorrect disparity values<sup>3,14,15</sup>. This process does not produce a clear estimate for ACRDS, since large peaks are predicted to occur at different locations at different spatial scales<sup>19</sup> (Fig. 3).

The perception of depth in ACRDS is further complicated by the existence of both first- and second-order mechanisms in stereoscopic processing. The discussion above considers the disparity information present in the Fourier components of the image, and how they might be combined. However, depth can also be perceived in second-order stimuli, containing informative disparities in variations in contrast, rather than in the underlying texture<sup>31</sup>. Evidence for the existence of second-order channels has been provided by both psychophysical experiments, showing that participants can perceive depth in these stimuli<sup>31–41</sup>, and physiological experiments, showing disparity-tuned responses to contrast envelopes<sup>14</sup>. Tanaka and Ohzawa<sup>14</sup> accounted for the responses of second-order neurons using a variation of the energy model. This model takes as its input not raw image values, but the outputs of monocular energy filters (Fig. 4). This monocular energy calculation is followed by a binocular energy calculation, with filters tuned to a much lower spatial frequency.

Wilcox and Hess<sup>40</sup> showed that depth can be perceived from contrast envelopes even when the underlying random noise patterns are completely uncorrelated, provided that there is an informative disparity in the contrast variation. Sensitivity to depth from second-order stimuli is much poorer than that for first-order stimuli<sup>41</sup>, and allows simple depth judgements to be made, but not the perception of 3D shape<sup>42</sup>. Second-order mechanisms will also provide disparity-tuned responses to reversed polarity stimuli, because the rectifying nonlinearity captures the magnitude, but not the contrast polarity, of luminance variations. An important distinction between first- and second-order channels is that the latter will signal depth in the forward, rather than reversed, direction. Cogan *et al.*<sup>18</sup> proposed that depth from ACRDS relies on second-order mechanisms. In their experiments, they found forward-depth perception for low density stimuli, but no reliable depth discrimination for high-density ACRDS.

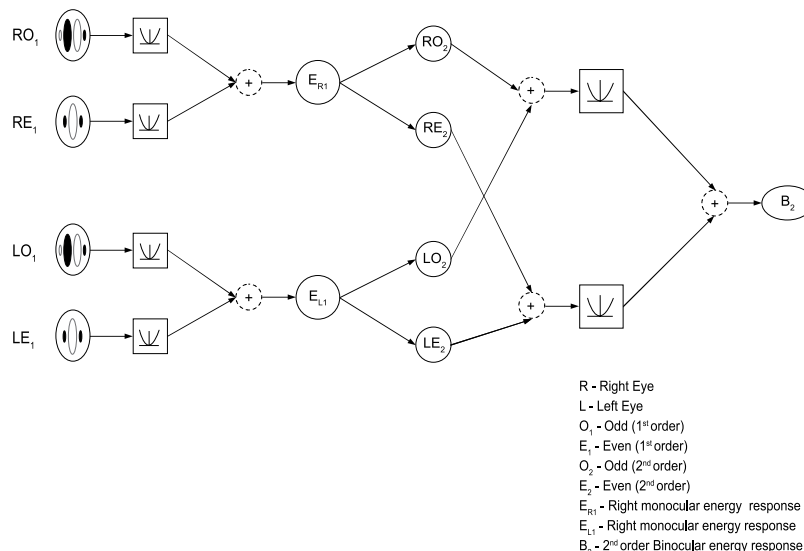
The perception of depth in both CRDS and ACRDS also depends strongly on the presence of features at different disparities in the stimulus, such that the depth of a target can be judged based on its disparity relative to that at other locations. Large changes in absolute disparity, when they are not accompanied by changes in relative disparity, can go unnoticed by participants<sup>43–45</sup>. Sensitivity to depth differences also falls as the spatial separation between target and reference increases<sup>46</sup>. Karmihirata *et al.*<sup>47</sup> argued that, due to the relatively weak disparity provided by ACRDS,



**Figure 3.** (a) An inverted disparity tuning function in response to an anti-correlated stereogram. (b–d) This produces peak responses that are offset from the tuning of the filter for correlated stimuli, which can signal the opposite sign of disparity to that of the neuron’s preferred disparity. The sign of this false peak depends on the tuning frequency of the filter and the stimulus disparity. Each plot shows the disparities of the peak responses to anti-correlated stimuli, for a filter that is tuned to the same disparity as the stimulus. Two peaks are plotted, one at a more crossed disparity, and one at a more uncrossed disparity, as indicated in (a). Responses are plotted as a function of disparity, averaged over 1000 Gaussian random noise stimuli. Individual plots show results for three different spatial frequencies. Depending on the combination of the stimulus disparity and the spatial frequency tuning of the filter, the two peaks either both signal forward-depth, or signal conflicting directions of depth.

depth is only perceived when a correlated reference is present, and there is no spatial gap between this and the anti-correlated target. Using stimuli in which the target was a circular region of anti-correlated dots, surrounded by an annulus of correlated dots, they showed that reversed-depth was perceived when there was no-gap, but that this deteriorated when a small-gap was presented. The lack of a gap means that, once a non-zero disparity is incorporated into the stimuli, there will be overlap between the dots in the target and surround, such that both correlated and anti-correlated dots will fall into the receptive fields of neurons aligned with the vertical edges of the stimulus. This creates regions of decorrelation at the edges of the stimuli, on a different side in each eye. Decorrelation of this type occurs naturally through half-occlusion, whereby parts of a stimulus are visible to one eye but not the other (Fig. 5)<sup>48–52</sup>. Depth is perceived in these stimuli, consistent with this geometric interpretation. It is thus possible that the perception of depth in this case relates to the presence of decorrelation, although it should be noted that this would not explain depth discrimination in other cases<sup>21</sup> in which there was a horizontal edge.

The first purpose of the current study was to understand the contribution of decorrelated edge regions, and the spatial separation between target and surround, on the perception of depth in ACRDS. This was done by using stimuli consisting of (i) an anti-correlated circular target surrounded by a correlated annulus (ii) a vertical edge between correlated and anti-correlated regions and (iii) a horizontal edge between correlated and anti-correlated regions. In the first two stimuli, there are regions containing both correlated and anti-correlated dots, whereas in the horizontal stimuli each region contains exclusively correlated or anti-correlated dots. The second purpose was to understand the contribution of mechanisms sensitive to first- and second-order disparities, and to monocular image regions, to the perception of depth in RDS. We did this firstly by modelling the responses of first- and second-order mechanisms to our stimuli, and secondly by manipulating the presence of decorrelated regions in our psychophysical experiments.



**Figure 4.** A second-order binocular energy model. The monocular energy responses are calculated separately for each eye. The first-order filtering responses are squared and summed, to produce monocular energy responses ( $E_{R1}$  and  $E_{L1}$ ). This forms the input to the second order filters where the pairs of monocular filters are summed, squared and combined to calculate the second-order binocular energy response.

## Psychophysical Experiment

**Methods.** *Participants.* 10 participants (8 females, mean(*std*) age 24.5(9.6)) completed the experiment. All had normal or corrected to normal vision, and stereoacuity of at least 50 arc sec, as measured using the Stereo Optical Butterfly Stereotest. All work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). The study procedures were approved by the University of Essex Ethics Committee (Application No. JA1601). All participants gave informed written consent and received payment for their participation.

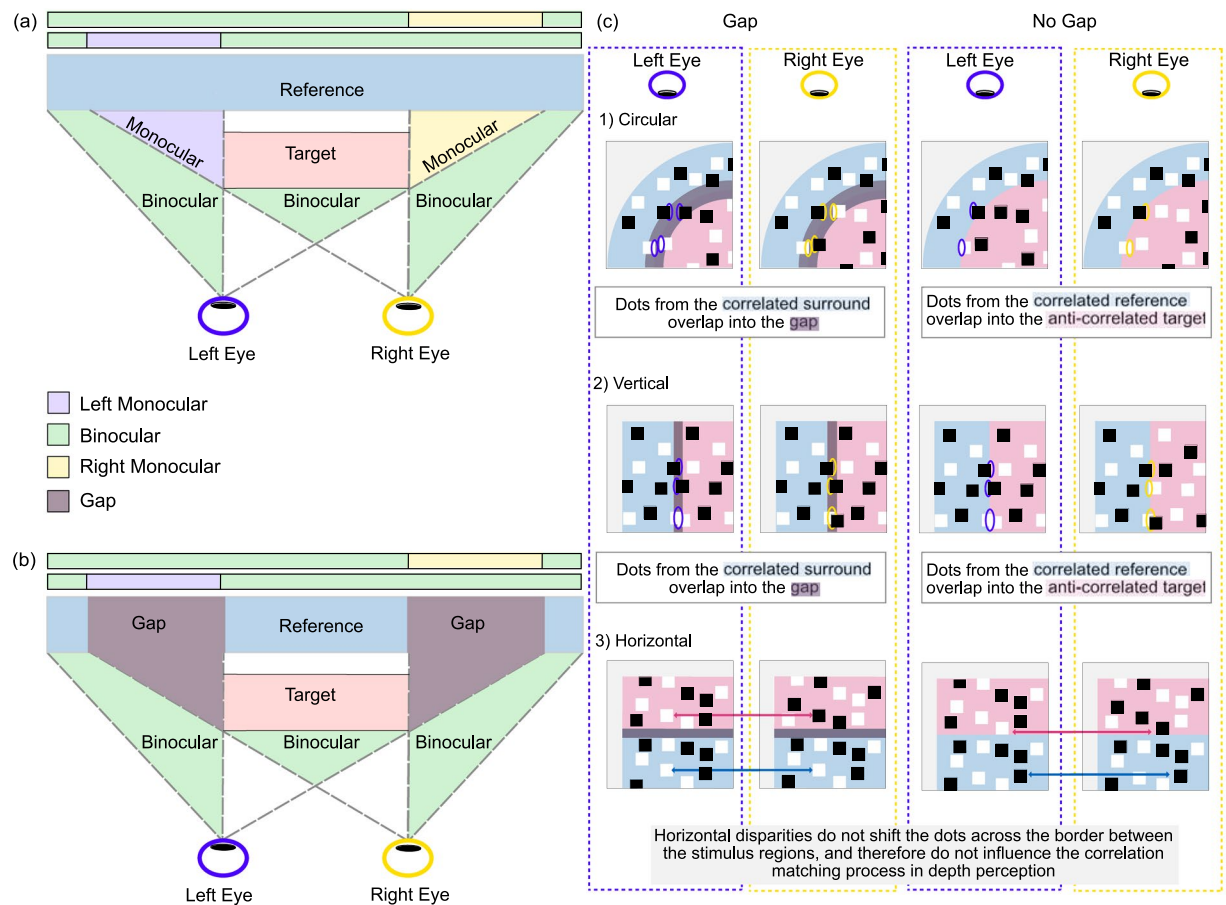
*Apparatus.* Stimuli were presented on a VIEWPIXX 3D monitor, viewed from a distance of 40 cm. The monitor screen was 52 cm wide and 29 cm tall. The screen resolution was 1920 by 1080 pixels, with a refresh rate of 120 Hz. Each pixel subtended 2.4 arc min. Stereoscopic presentation was achieved using a 3DPixx IR emitter and NVIDIA 3D Vision LCD shutter glasses. The cross-talk between the left and right images, measured using a Minolta LS-110 photometer, was 0.12%. Participants' responses were recorded using the computer keyboard. Stimuli were generated and presented using MATLAB and the Psychophysics Toolbox extensions<sup>53–55</sup>.

*Stimuli.* Stimuli in all conditions were random dot stereograms, consisting of 12 arc min square red ( $27.0 \text{ cdm}^{-2}$ ) and black ( $0 \text{ cdm}^{-2}$ ) dots against a red ( $13.5 \text{ cdm}^{-2}$ ) background. Equal numbers of red and black dots were presented, with a total density of  $1.12 \text{ dots/degree}^{-2}$ . In all cases, stimuli consisted of a correlated reference region, presented with 0 disparity, and a test region which was either correlated or anti-correlated. The dots were in all cases randomly repositioned on each frame. The test region was presented with disparities of  $\pm 5.5$ ,  $\pm 11$  and  $\pm 22$  arc min. Stimuli were presented in three conditions: circular, horizontal and vertical (Fig. 6(a–i)).

*Circular.* The target was a circular region with a diameter of 4.4 degrees. The reference was a surrounding annulus, with no-gap, a small-gap (19.2 arc min) or a large-gap (38.4 arc min) between target and the inside of the reference region (Fig. 6(a–c)). The inner diameter of the reference annulus depended on the size of the gap between the test and reference, and the outer diameter was 1.83 degrees larger than the inner radius. The circular stimuli were presented with the centre of the test region 5.5 degrees below the fixation cross.

*Horizontal Edge.* The reference and test regions were both a rectangle with a width of 5.5 degrees and a height of 2.75 degrees. Both were centred 5.5 degrees to the right of fixation, with the test presented above and the reference presented below the fixation cross. There were three levels of vertical gap; no-gap, a small-gap (19.2 arc min) or a big-gap (38.4 arc min) between the reference and test regions (Fig. 6(d–f)).

*Vertical Edge.* The reference and test regions were both a rectangle with a width of 2.75 degrees and a height of 5.5 degrees. Both were centred 5.5 degrees below the fixation cross, with the test presented to the right and the reference presented to the left. As with the previous conditions, there was a horizontal separation of either no-gap, a small-gap (19.2 arc min) or a big-gap (38.4 arc min) between the reference and test regions (Fig. 6(g–i)).



**Figure 5.** (a) Da Vinci stereopsis: when an object (Target) occludes different parts of a background (Reference) from each eye's perspective, there is a difference between the images in the left and right eyes<sup>50</sup>. However, (b) Including a gap between the Target and the Reference, reduces or eliminates this monocular occlusion. (c) A schematic representation of the process of Da Vinci stereopsis for different types of edges. The different image for each eye can contain matched (correlated) and unmatched (uncorrelated) information. With the exception of the horizontal condition (3), stimuli without a gap between the target and the reference (such as 1 and 2) create a potential area of overlap between correlated and uncorrelated information.

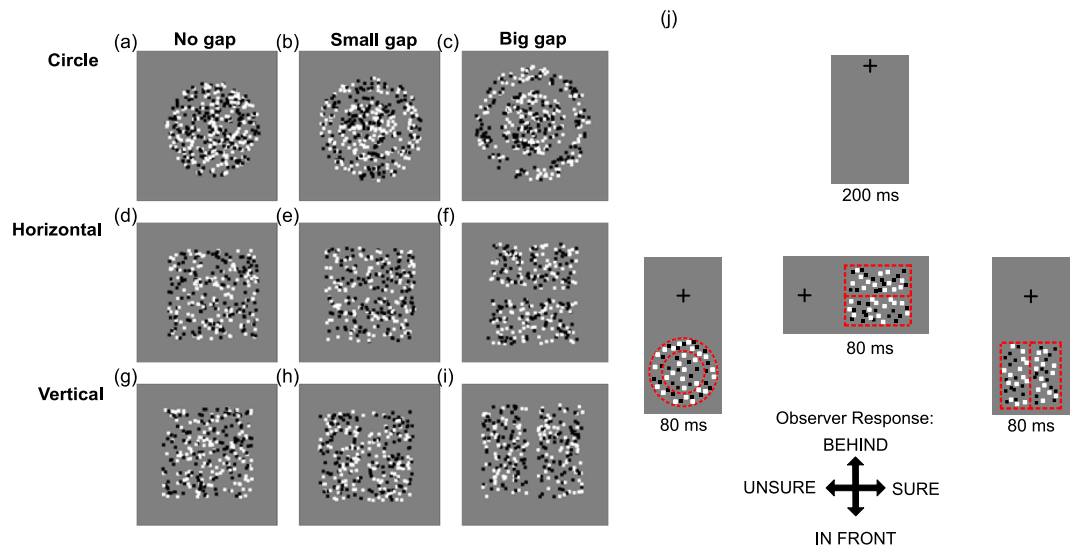
**Procedure.** Each trial began with the presentation of a central fixation cross for 200 ms, followed by the presentation of a stimulus for 80 ms. The fixation cross remained visible throughout each trial block.

After stimulus presentation, the participant was required to make two responses. The first was to indicate whether the target appeared closer (down arrow) or further away (up arrow) than the reference. The second was to indicate whether they felt confident in their response (right arrow) or that they were guessing (left arrow). The next trial began after the two responses had been made (Fig. 6(j)).

Each participant completed 18 blocks of trials, for all combinations of the three configurations, three separations, and two correlation conditions. In each block, each of the 6 disparities was presented 20 times. Blocks were presented in a randomised order, and separated over two or more sessions.

**Results. Depth Perception.** To determine whether participants' judgements shifted from 'far' to 'near' (or vice versa) as disparity shifted from uncrossed to crossed, results were analysed with a generalised linear mixed effects model. In this model, disparity was used as a predictor, with a probit linking function, and random slopes and intercepts across participants were included. A separate model was fit for each combination of shape, correlation and stimulus separation. This model allows us to determine whether, at the population level, forward- or reversed-depth was perceived, while also taking account of variation across participants.

Previous studies have found significant individual differences in depth perception in ACRDS, for example with some participants perceiving reversed-depth but others perceiving no depth<sup>21</sup>. We therefore fit the same generalised linear model, with a probit linking function, to the data for each participant separately. In each case, we then determined whether that participant perceived forward- or reversed-depth for each stimulus based on whether the slope parameter of the model had a significant positive or negative value. The results are shown in Fig. 8, which plots the number of participants perceiving forward-depth, reversed-depth, or no significant depth, for each stimulus type.



**Figure 6.** (a) Schematic of stimuli used for each condition. (a) For the circular condition, the target was always the inside circle. (b) For the horizontal condition the target was always the top rectangle and (c) for the vertical condition the target was always the right rectangle. The task in each case was to identify whether the target was behind or in front of the reference. Participants were required to stay fixated on the cross at all times. (b) Participants were required to indicate for each presentation whether the target was in front of or behind the reference, immediately after which a confidence decision (sure or unsure) was required. The procedural schematic shows a border to separate target and reference, and the relative location of the fixation cross for information only.

**Circular Stimuli.** The proportion of near judgements is plotted as a function of disparity. Results are plotted separately for CRDS and ACRDS and for the three separation distances. Figure 7(a–c) shows the mean results across participants. For CRDS, there was a positive slope for all separations indicating forward-depth. In contrast there was no significant slope in either direction for ACRDS (Table 1).

Individual analyses (Fig. 8) showed that the majority of participants perceived forward-depth for CRDS, while some were not able to reliably report the direction of depth for the larger stimulus separations. For CRDS, in most cases there was no reliable perception of depth. With no separation, 4/10 participants reported reversed-depth, but no reversed-depth was perceived for larger separations, consistent with recent findings<sup>47</sup>. However, for the largest separation, there was evidence of reliable perceived depth in the forward direction for two participants.

**Horizontal Stimuli.** Data are presented in Fig. 7(d–f) in the same format as for circular stimuli. For CRDS, there was a positive slope for all separations. For ACRDS, there was a significant positive slope for the small separation, consistent with forward, rather than reversed, depth (Table 1).

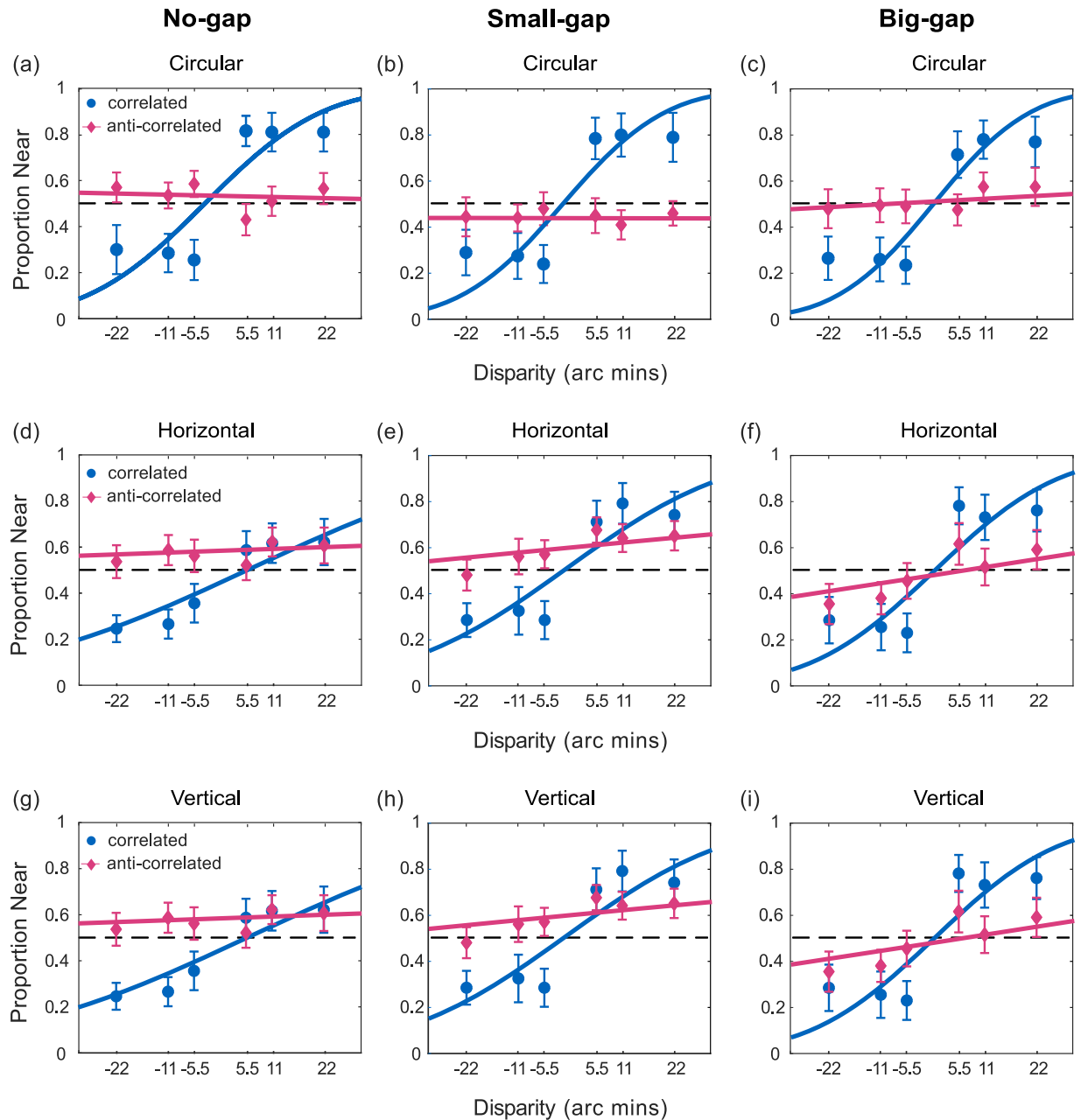
Individual analyses (Fig. 8) showed that the majority of participants perceived forward-depth for CRDS, although this proportion decreased with increasing separation. For ACRDS, one participant perceived depth in the reverse direction when there was no-gap, and a minority of participants perceived depth in the forward direction.

**Vertical Stimuli.** Data for the vertical stimuli are presented in Fig. 7(g–i). For CRDS, there is a positive slope for all separations. For ACRDS, there was a significant positive slope for the two non-zero gaps, consistent with forward-depth (Table 1).

Individual analyses (Fig. 8) showed very similar results to those found with a horizontal edge. The majority of participants perceived forward-depth for CRDS, and this proportion decreased with increasing separation. For ACRDS, the majority of participants did not reliably perceive depth, but a minority did consistently perceive forward-depth, and there was only one example of reliable reversed-depth.

**Confidence.** Mean confidence ratings are plotted as a function of disparity for each combination of shape, correlation and distance in Fig. 10. Results are averaged across signs of disparity since, unlike depth judgements, we do not expect opposite results for near and far stimuli.

These results were analysed using a linear mixed effects model, with the shape (circular, horizontal or vertical) and correlation (correlated or anti-correlated) as categorical factors, and separation and disparity as linear covariates. A correlation-by-distance interaction term was also included to determine whether separation had a greater effect on confidence for ACRDS than for CRDS. Participant was included as a random factor, with random intercepts and slopes against disparity. The results are summarised in Table 2. Confidence ratings were significantly lower for ACRDS than for CRDS. They were also significantly lower for the horizontal and vertical

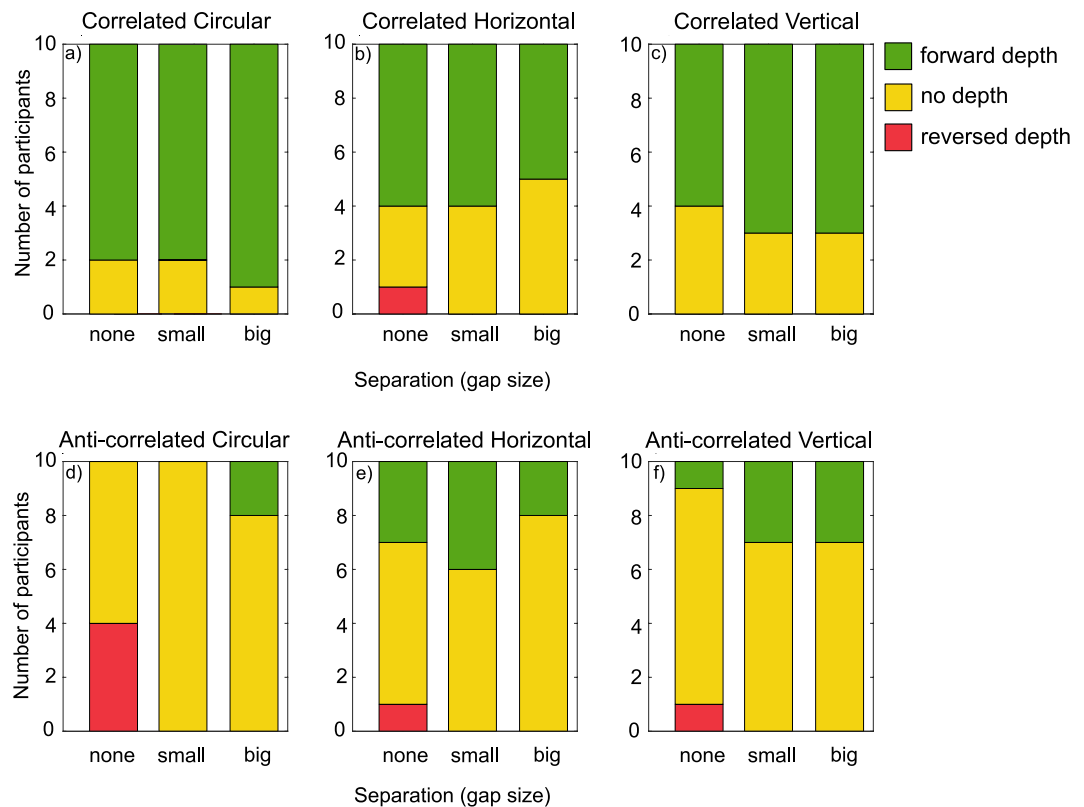


**Figure 7.** Mean proportion of near responses as a function of disparity for stimuli with no-gap (left), small-gap (middle) and big-gap (right). Circular results are shown in (a–c), Horizontal results in (d–f) and Vertical results in (g–i). Separate plots chart the responses to anti-correlated (pink) and correlated (blue) stimuli. The dashed line for each plot reflects performance at chance, uncrossed and crossed disparities are shown with negative and positive values. Error bars are  $\pm 1$  standard error of the mean.

stimuli than for the circular stimuli. Ratings tended to increase with increasing gap size, but were not affected by disparity. There was no significant distance-by-correlation interaction, meaning that the separation did not affect confidence differently for CRDS and ACRDS. The effects of correlation, shape and distance are summarised along with the mean ratings in Fig. 9.

**Relationship between Confidence and Performance.** For each stimulus, we recorded a near/far depth judgement and an indication of whether or not the participant felt confident in this judgement. Figure 10 plots heatmaps of the relationship between confidence and performance, for CRDS and ACRDS, summed over all participants, disparities, separations and stimulus types. Performance was summarised as the number of consistent responses such that, for example, 20 near responses out of 20, and 20 far responses, were both coded as 100% consistent, while a stimulus for which there were 10 near and 10 far responses was coded as 50% consistent. This consistency measure does not depend on whether responses are in the forward or reversed direction, relative to the stimulus





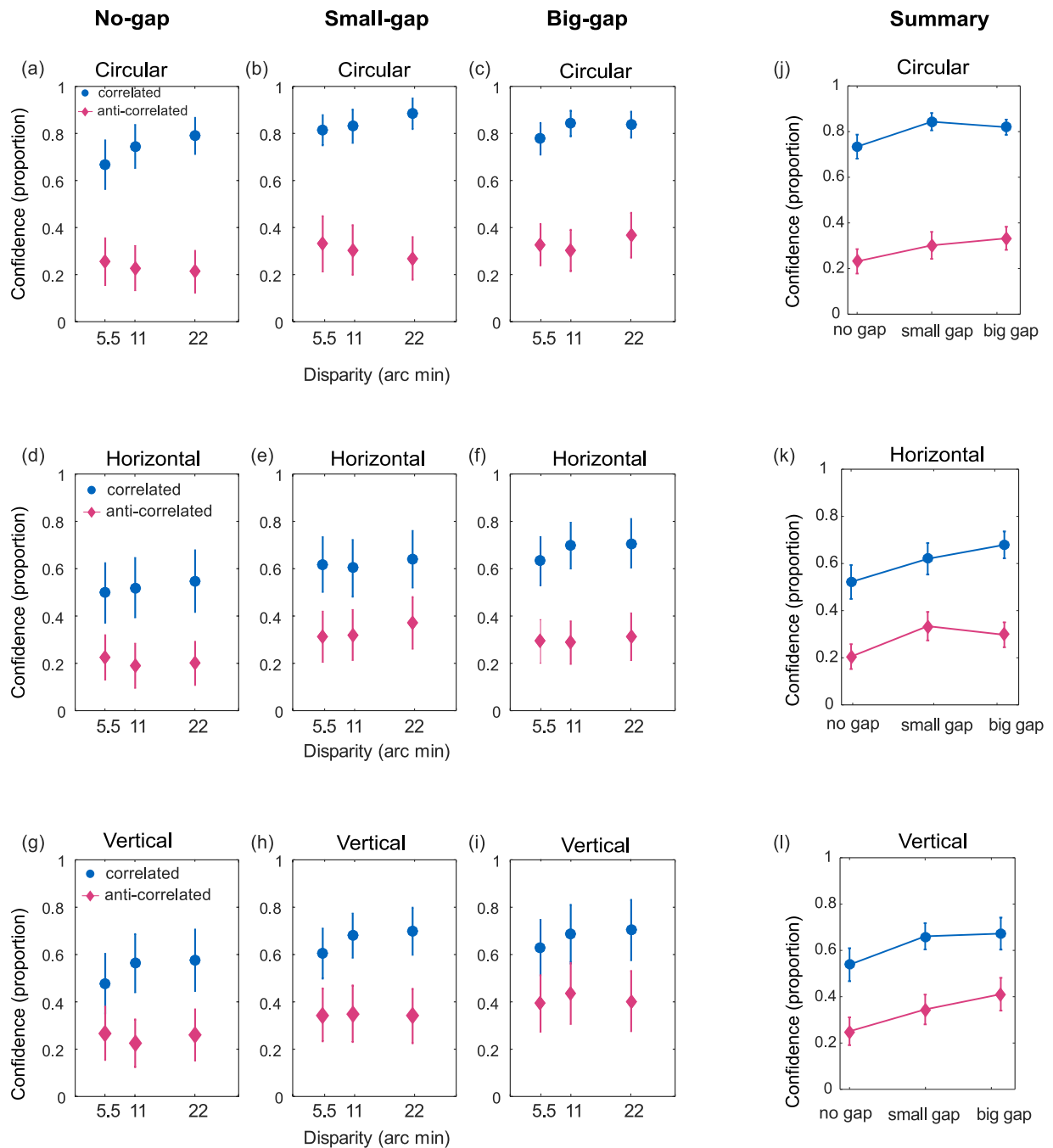
**Figure 8.** The number of participants reporting forward (green) reversed (red) or no (yellow) depth.

Shape	Correlation	Gap Size	$\beta$	SE	tStat	DF	p Value
Circular	CRDS	0	0.05095	0.018469	2.7586	58	0.0077524*
Circular	CRDS	19.2	0.058554	0.020582	2.845	58	0.0061278*
Circular	CRDS	38.4	0.061829	0.021556	2.8682	58	0.0057471*
Circular	ACRDS	0	-0.0011257	0.0021927	-0.5134	58	0.60962
Circular	ACRDS	19.2	-7.8727e-05	0.0013605	-0.057867	58	0.95405
Circular	ACRDS	38.4	0.0027754	0.0021185	1.3101	58	0.19532
Horizontal	CRDS	0	0.042448	0.015851	2.6779	58	0.0096197*
Horizontal	CRDS	19.2	0.038207	0.015074	2.5347	58	0.013976*
Horizontal	CRDS	38.4	0.039287	0.015882	2.4736	58	0.016324*
Horizontal	ACRDS	0	0.004145	0.0030626	1.3534	58	0.18117
Horizontal	ACRDS	19.2	0.0069563	0.0029483	2.3594	58	0.021692*
Horizontal	ACRDS	38.4	0.0045466	0.0033093	1.3739	58	0.17476
Vertical	CRDS	0	0.023867	0.011142	2.1421	58	0.036395*
Vertical	CRDS	19.2	0.03677	0.014867	2.4733	58	0.016339*
Vertical	CRDS	38.4	0.048554	0.017404	2.7899	58	0.0071229*
Vertical	ACRDS	0	0.0018335	0.0018846	0.97288	58	0.33465
Vertical	ACRDS	19.2	0.0050179	0.0022086	2.2719	58	0.026817*
Vertical	ACRDS	38.4	0.007975	0.0036747	2.1702	58	0.034096*

**Table 1.** Estimates of effects from the generalised linear mixed effects regression for the proportion of near versus far responses in CRDS and ACRDS stimuli. \*Indicates significant p value.

disparity. For CRDS, results are clustered in the top-right corner, indicating generally high levels of both consistency and confidence. For ACRDS, confidence was generally low, consistent with a low level of consistency in responses.

*Discussion.* The aim of our psychophysical experiment was to understand how the spatial separation (the size of the gap) and overlapping correlated and anti-correlated elements influence our perception of depth in ACRDS.

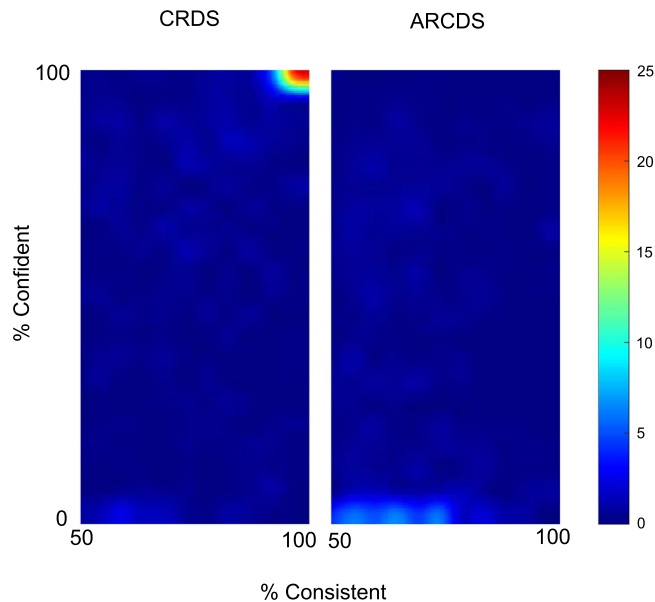


**Figure 9.** (a–i) Mean confidence (proportion) for CRDS (blue) and ACRDS (pink) are plotted as a function of disparity and gap size and shape. Error bars show  $\pm 1$  standard error of the mean. (j–l) Summary of confidence in depth detection with each shape on separate plots. Plotted as a function of gap size for CRDS (blue) and ACRDS (pink), an averaged over disparity.

Predictably, our results show forward-depth perception for all CRDS stimuli. Depth perception for horizontal and vertical stimuli was not influenced by the size of the gap or type of stimulus. In addition for CRDS, confidence increased with gap size. Confidence was highest for circular stimuli and lower for horizontal and vertical stimuli.

In contrast, for ACRDS confidence in judgements was at or below 50% for all conditions. These are significantly lower than the respective confidence ratings for CRDS. Furthermore, depth perception for ACRDS was variable depending on the type of stimulus.

For circular ACRDS there was, on average, no depth perception for any gap size. However, reversed-depth as predicted by Aoki *et al.*<sup>17</sup> was seen by some participants in the no-gap condition. For the small-gap, there was no depth discrimination shown by any participants, and for the big-gap only two participants reported forward-depth.



**Figure 10.** Relationship between consistency of, and confidence in, responses, for CRDS (left) and ACRDS (right). Colour indicates the percentage of responses, for each trial, at each level of consistency and confidence, averaged across all participants and stimuli. High levels of both consistency and confidence were recorded for CRDS, and low levels of both measures for ACRDS.

Shape	$\beta$	SE	tStat	DF	p Value	CI	CI
Circular	23.162	3.7038	6.2536	533	8.2323e-10	15.886	30.438
Horizontal	-4.0333	1.0966	-3.678	533	0.00025894*	-6.1876	-1.8791
Vertical	-2.5611	1.0966	-2.3355	533	0.019889*	-4.7153	-0.40689
Correlation	-14.711	2.369	-6.2099	533	1.0675e-09*	-19.365	-10.057
Gap Size	2.5111	0.93709	2.6797	533	0.0075966*	0.67028	4.3519
Disparity	0.083911	0.065264	1.2857	533	0.1991	-0.044296	0.21212
Correlation x Gap	-0.15556	1.0966	-0.14185	533	0.88725	-2.3098	1.9987

**Table 2.** Estimates of effects from the linear mixed effects regression for confidence judgements in CRDS and ACRDS stimuli. \*Indicates significant p value.

For horizontal stimuli, on average there was a tendency towards forward-depth for ACRDS with a small-gap, but no forward or reverse perception of depth for stimuli with no-gap or a big-gap. While the majority of participants reported no depth, a minority reported forward-depth, and one participant reported reversed-depth.

Finally, for vertical stimuli, on average depth was perceived in the forward direction for the small- and large-gap ACRDS, while there was no perception of depth with the no-gap stimuli. This trend reflected the forward-depth perception in a minority of participants. In both the horizontal and vertical edge conditions, perception of forward-depth in CRDS conditions was lower than in the circular conditions. This may be as a result of the surround annulus in the circular condition providing more points of reference (zero disparity), allowing for better judgement of depth for the central target. While it may be of interest to test this edge perception with an inner and outer rectangle, this was not the purpose of our investigation. However, we would predict that the complexity provided by such a target would provide the same results (for CRDS and ACRDS) as the circular conditions. We will discuss this prediction in more detail in the general discussion.

Aoki *et al.*<sup>17</sup> found that removing the gap between the target and the surround for ACRDS increased the perception of reversed-depth. However the removal of the gap results in the potential for overlap between the anti-correlated dots in the target and the correlated dots in the surround. This results in decorrelation (an average correlation of zero) at the edges on opposite sides of the target in each eye. This occlusion occurs naturally in the phenomenon of da Vinci stereopsis when the edges of vertically oriented stimuli occlude the surface from the opposite eye<sup>50</sup> as illustrated in Fig. 5(a–c). If these uncorrelated regions were taken to indicate the presence of half-occlusions, we would predict that depth would be perceived in the reverse direction.

Our results indicate that there was little evidence of reversed-depth for the vertical edges, and one condition (small-gap) showed a significant tendency for depth to be perceived in the forward direction. There was only one participant who reported reversed-depth for the horizontal no-gap condition.

Our results are consistent with the findings of Aoki *et al.*<sup>17</sup>, where removing the gap did result in reversed-depth for some observers. However, this only occurred for the circular condition, and no reliable perception of

reversed-depth was reported for the horizontal or vertical no-gap conditions. This suggests that the complexity of the circular stimulus (particularly when there is no separation between the target and its zero-disparity reference) is more likely responsible for the perception of reversed-depth, than is the removal of the gap.

This study found no robust evidence for reliable reversed-depth perception in any condition, supporting the findings by Hibbard *et al.*<sup>19</sup>. The anti-correlated circular condition with no-gap produced the highest proportion of reversed-depth perception, consistent with previous results<sup>17,47</sup>. Furthermore the circular no-gap and small-gap conditions were the only conditions not reporting any forward-depth perception. The trend towards depth perception in the forward direction, while not consistent with the information provided by first-order disparity channels, is predicted by the depth signalled by second-order mechanisms. The conflicting depth signalled by first- and second-order mechanisms for ACRDS is explored in the following section.

### First- and Second-Order Responses to Correlated and Anti-Correlated Random Dot Stereograms

The perception of depth depends on multiple mechanisms, including first- and second-order channels, tuned to a variety of scales and orientations. How these contribute to the perception of depth depends on how the information they provide is processed in higher cortical areas. In order to understand the contribution of first- and second-order mechanisms to perceived depth in CRDS and ACRDS, we modelled the responses of these mechanisms to our stimuli.

**Methods.** *First- and Second-Order Mechanisms.* First-order disparity-sensitive mechanisms were implemented using a standard binocular energy model. The first stage of this model was a quadrature pair of Gabor filters:

$$\begin{aligned} G_{E1} &= \exp\left(-\frac{x^2}{2\sigma_H^2} - \frac{y^2}{2\sigma_V^2}\right) \cdot \cos(2\pi fx) \\ G_{O1} &= \exp\left(-\frac{x^2}{2\sigma_H^2} - \frac{y^2}{2\sigma_V^2}\right) \cdot \sin(2\pi fx) \end{aligned} \quad (1)$$

where  $(x, y)$  is the spatial position,  $f$  is the spatial frequency and  $\sigma_H$  and  $\sigma_V$  determine the horizontal and vertical extent of the Gaussian envelope. All filters were tuned to a vertical orientation, and spatial frequencies of 4, 8 and 16 cycles/degree were used. The values of  $\sigma_H$  and  $\sigma_V$  depended on spatial frequency as follows:

$$\begin{aligned} \sigma_H &= \frac{0.39}{f} \\ \sigma_V &= \frac{0.78}{f} \end{aligned} \quad (2)$$

The left and right images were both convolved with each filter, to produce four responses:

$$\begin{aligned} L_{E1} &= L * G_{E1} \\ L_{O1} &= L * G_{O1} \\ R_{E1} &= R * G_{E1} \\ R_{O1} &= R * G_{O1} \end{aligned} \quad (3)$$

Binocular energy was calculated as:

$$B_1 = (L_{E1} + R_{E1})^2 + (L_{O1} + R_{O1})^2 \quad (4)$$

The second-order mechanisms included two filtering stages. The first used the same filters as the first-order mechanisms. Left and right monocular energy responses were calculated as:

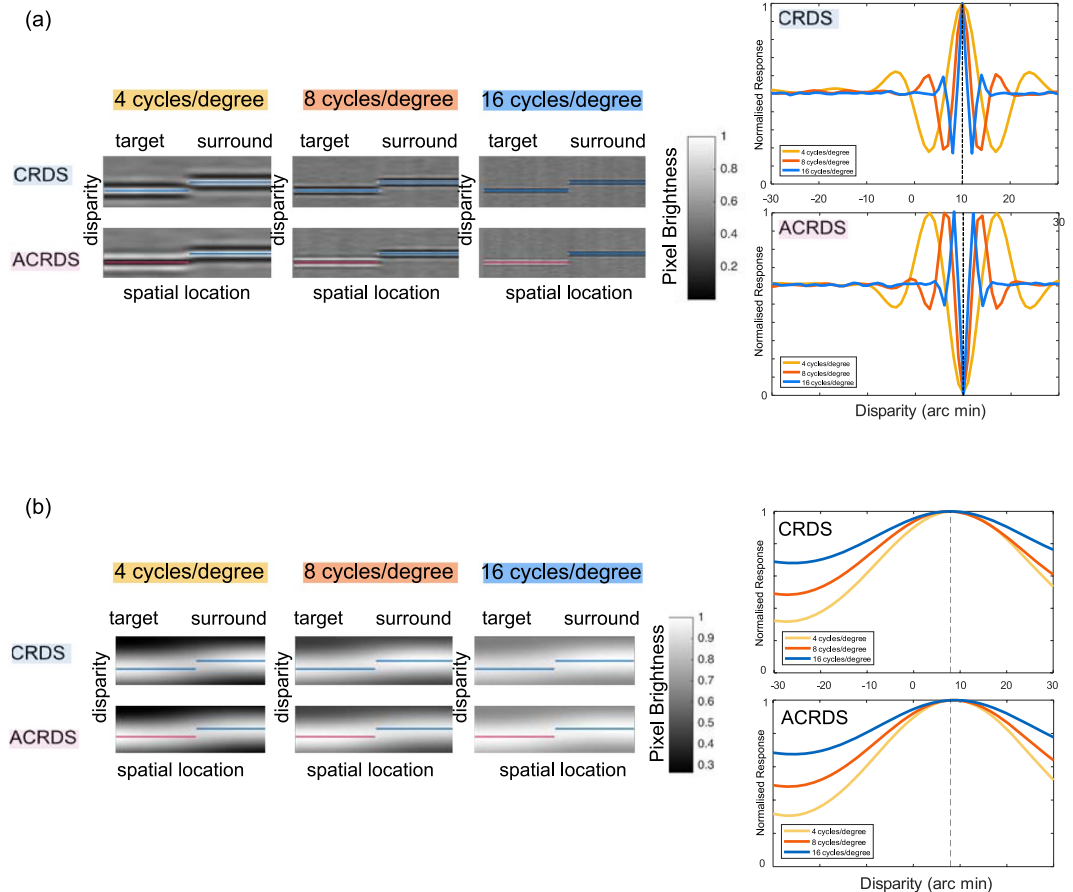
$$\begin{aligned} E_{L1} &= L_{E1}^2 + L_{O1}^2 \\ E_{R1} &= R_{E1}^2 + R_{O1}^2 \end{aligned} \quad (5)$$

Second-order filters ( $G_{E2}, G_{O2}$ ) had the same shape as first-order filters, but a frequency tuning of 0.8 cycles/degree. Monocular responses were then calculated by convolving the left and right energy responses with the second-order filters,

$$\begin{aligned} L_{E2} &= E_{L1} * G_{E2} \\ L_{O2} &= E_{L1} * G_{O2} \\ R_{E2} &= E_{R1} * G_{E2} \\ R_{O2} &= E_{R1} * G_{O2} \end{aligned} \quad (6)$$

and combining these to produce the second-order energy response:

$$B_2 = (L_{E2} + R_{E2})^2 + (L_{O2} + R_{O2})^2 \quad (7)$$



**Figure 11.** (a) The responses of first-order mechanisms, tuned to 4, 8 and 16 cycles/degree, to CRDS (top row) and ACRDS (bottom row). In each image, horizontal and vertical position represent the spatial location and preferred disparity of the filters, respectively, and pixel brightness the strength of response. The target (correlated or anticorrelated) is on the left of each image and the zero-disparity surround is on the right. The disparities of these are indicated by the horizontal lines (blue for correlated regions, pink for anti-correlated regions). The two graphs on the right show the disparity tuning, averaged over the target regions, for all three frequencies, for CRDS (top) and ACRDS (bottom). (b) The responses of second-order mechanisms, with first-stage filters tuned to 4, 8 and 16 cycles/degree, to CRDS (top row) and ACRDS (bottom row). The second stage filters are in all cases tuned to 0.8 cycles/degree.

**Stimuli.** Firstly, we calculated the first- and second-order binocular energy responses for stimuli with a simple disparity difference. Responses to stimuli with the same random dot pattern as used in the psychophysical experiments were analysed. In each stimulus, the upper half was correlated and presented with zero disparity, and the lower half either correlated or anti-correlated and presented with a disparity of 10 arc min.

Secondly, we calculated responses, as a function of spatial position, to the actual horizontal, vertical and circular stimuli used in the psychophysical experiments.

**Procedure.** For the first analysis, we calculated the first- and second-order binocular energy responses for 1000 samples of the stimuli with a simple disparity step difference. In all cases, the values used to calculate the energy response were taken from the central row of the left image, and from a row in the right image corresponding to a horizontal disparities ranging between  $\pm 30$  arc min, sampled at intervals of 1 arc min. All convolutions were calculated by multiplication in the Fourier frequency domain.

For the second analysis, we calculated first- and second-order binocular energy responses for 20 simulated trials, using filters tuned to 10 arc min crossed, 0 arc min, and 10 arc min uncrossed disparities, for a stimulus with a 10 arc min crossed disparity. This allowed us to calculate the energy response, as a function of spatial position, from filters tuned to the zero disparity reference, and to the forward and reversed directions of the target disparity.

**Results.** Figure 11(a,b) display the results of the first analysis, where (a) shows the mean, over 1000 trials, of the first order energy response, as a function of vertical position and horizontal disparity. Results are shown separately for each of the three spatial frequencies and are plotted separately for CRDS and ACRDS. Results were normalised for each frequency by dividing by the maximum of the mean response. The horizontal lines show the target- and surround regions, with a non-zero and zero disparity, for correlated (blue) and anticorrelated (pink)

regions. For CRDS, there is a peak in response for all frequencies at the correct disparity, flanked by a minimum in response, the disparity of which depends on the frequency tuning of the filter. This pattern reflects the tuning function shown in Fig. 1. For the ACRDS, this response is reversed, such that there is a minimum response at the correct disparity. Reversed-depth has been attributed to evidence against the presence of a stimulus at the correct disparity<sup>17</sup>, or to the presence of the flanking false peaks<sup>21</sup>. Figure 11(b) shows the mean of the second-order energy response, plotted in the same way as for the first-order response. For CRDS, these results show a very similar pattern to the first-order response, except with much coarser spatial resolution, due to the low-frequency tuning of the second-order filters. For the ACRDS, the inversion of the disparity tuning function is absent. This means that, for ACRDS, second-order filters provide a response that is consistent with the forward disparity, and consistent across the three spatial frequencies of first-order filtering.

The results of the second analysis (Fig. 12) shows the mean energy responses, over the 20 simulated trials and three spatial frequencies, of the first- and second-order binocular energy responses, at the three different disparities for CRDS and ACRDS for each of the shape and gap conditions. The strongest response for the correlated zero-disparity reference (surround for circular, bottom for horizontal, and left for vertical) region occurs for the filters tuned to zero disparity for both CRDS and ACRDS in all shape and gap conditions for both first- and second-order mechanisms. For CRDS, the strongest response for the target (centre for circular, top for horizontal, and right for vertical) region signals forward-depth for both first- and second-order mechanisms in every shape by gap condition. In contrast, the strongest responses for the target (centre for circular, top for horizontal, and right for vertical) in ACRDS occurs for the filter tuned to a reversed-depth direction for first-order mechanisms and forward-depth for second-order mechanisms. These results are consistent with previous discussions.

Our analysis shows that, for correlated stimuli, both first- and second-order mechanisms signal depth in the forward direction, for all shapes and gap sizes. For anti-correlated stimuli, the first-order channel signals reversed-depth while the second-order channel signals forward-depth.

**Discussion.** The purpose of the second study was to investigate first- and second-order contributions to the perception of depth. This analysis focused on modelling responses to the stimuli used in this study. The reversal of apparent depth found for some ACRDS has been linked to the inversion of the disparity tuning function of binocular cortical cells. This inversion of disparity tuning does not occur for second-order mechanisms, due to the rectifying non-linearity that occurs between the first- and second-stage filtering. This means that second-order units indicate depth in the forward direction for both CRDS and ACRDS.

Hibbard *et al.*<sup>15</sup> proposed that second-order mechanisms may provide some error correction to the first-order disparity estimations where regions are decorrelated<sup>15</sup>. This is consistent with the perception of forward-depth in ACRDS shown here, despite the lack of a clear binocular cross-correlation peak at the target disparity. While the first- and second-order responses have been shown to be highly correlated in natural stimuli, second-order mechanisms can increase the accuracy of depth perception, acting as a ‘back-up’ mechanism when first-order mechanisms do not provide a reliable signal<sup>31</sup>.

The existence of both first- and second-order mechanisms substantially complicates the attempt to link depth perception in ACRDS to neural population responses. It may also account for why there is so much variability across individual participants<sup>15</sup>. As well as providing an impoverished, inconsistent signal across scales in the first-order channels, conflicting estimates can arise between first- and second-order channels.

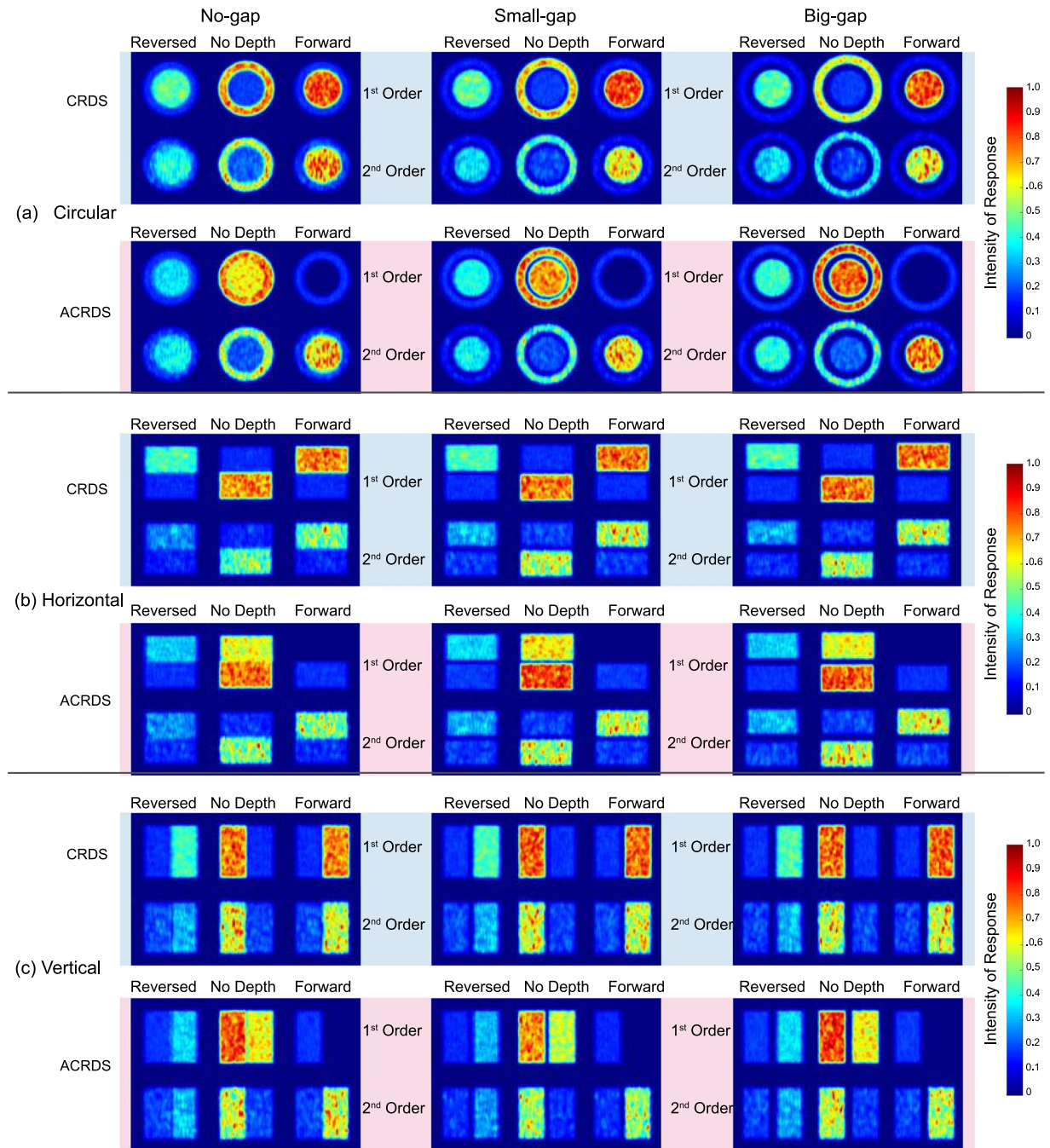
## General Discussion

Our findings suggest that neither particular spatial configurations of stimuli, nor gap sizes, contributed to robust perception of reversed-depth in ACRDS. As expected, CRDS stimuli were all perceived with depth in the geometrically predicted forward direction. For ACRDS, most participants reported no depth at all, and in some cases there were significant positive slopes consistent with forward-depth perception for the vertical (small-gap & big-gap) and horizontal (small-gap only) conditions. For the anti-correlated circular conditions, overall, there was very little perception of forward-depth for any gap size, with the exception of 2 participants when the gap was at its biggest. For the no-gap condition reversed-depth was reported by 4 participants.

In order to better understand the contributions of the first- and second-order disparity processing stages to the perception of depth, we modelled the responses to the stimuli used in this experiment using a standard binocular energy model, adapted to include second-order filters<sup>23</sup>. Results from the first-order energy response show a peak for CRDS at the correct disparity, as would be expected. For ACRDS, this peak is reversed. The second-order energy response shows a similar peak location as the first-order response for CRDS, but with much broader disparity tuning. However, for the ACRDS there was no inverse peak, as predicted by the rectifying non-linearity occurring prior to the second stage of filtering. This indicates that second-order mechanisms signal forward-depth for both CRDS and ACRDS.

The conflicting signals from the first- and second-order channels may explain the contradictory results between studies<sup>17,19</sup> and between our conditions, and participants. Wilcox and Allison<sup>31</sup> proposed that the second-order mechanism provides a back up to the first-order mechanism. When there is a reliable luminance-based disparity signal then the first-order mechanism will be relied upon. However, when there is an ambiguous signal, such as when using anti-correlated or uncorrelated dots, second-order signals are used<sup>31</sup>. Strong evidence to support this is found in individuals with strabismus<sup>56</sup>, a condition where the eyes are abnormally aligned. Standard clinical tests suggest that individuals with strabismus are stereo-deficient, however McColl *et al.*<sup>56</sup> demonstrated that some individuals were still able to perceive depth in their task. Wilcox and Allison<sup>31</sup> suggest this depth perception was via coarse disparity signals from the second-order mechanisms.

The specific contributions of the first- and second-order mechanisms to depth perception have been described as quantitative and qualitative, respectively<sup>31</sup>. Like motion, second-order stereopsis provides low-resolution depth



**Figure 12.** Heatmaps for all nine conditions of the simulation for our specific stimuli. Circular stimuli are on the top (a), horizontal in the centre (b) and vertical at the bottom (c). Each shape has rows for CRDS (top) and ACRDS (bottom), and the gaps in columns, with no-gap (left), small-gap (middle) and big-gap (right). The reversed column shows results for filters tuned to an uncrossed disparity of 10 arc min, the no-depth column for filters tuned to zero disparity, and the forward column for filters tuned to a crossed disparity of 10 arc min. Responses were summed over the three spatial frequencies and 20 trials. Results were normalised separately for the first- and second-order channels by dividing by the maximum mean response. In all cases, the correlated zero-disparity reference produces a strong response in the zero-disparity-tuned filters. For the first-order channel, a stronger response to the target is evident in the forward (CRDS) or reverse (ACRDS) direction. In contrast, for the second-order channel a stronger response to the target is evident in the forward direction for both CRDS and ACRDS.

estimates over a large range of disparities, where first-order mechanisms provide a high-resolution estimate of depth when the target is matched (correlated) in both eyes. Investigating the contribution of the second-order mechanisms, Ziegler and Hess<sup>42</sup> used an assortment of stimuli specifically designed to dissociate between linear and nonlinear mechanisms. The task required observers to identify the shape of a target. During testing the

display could switch between matched (correlated) to unmatched (uncorrelated) versions of the stimuli. All the observers performed normally and identified shape in matched mode, and as such a staircase converged on their identification threshold. However, when switching to unmatched stimuli, observers could no longer identify the shape, the staircase failed to converge and performance was at chance. Ziegler and Hess<sup>42</sup> found that while depth perception was demonstrated in all conditions, no reliable perception of shape was reported for any nonlinear (unmatched/uncorrelated) stimuli.

For depth perception, the first-order signal is often the default signal, switching to second-order signals when the first-order is ambiguous<sup>31</sup>. This would result in a prediction of forward-depth for all ACRDS conditions. However, if the second-order stimulus (for example as seen with no-gap or small-gap) is also ambiguous then potentially there is no useful depth information from either. Consider the heatmaps from our simulations, specifically those for the second-order responses to circular ACRDS (Fig. 12(a) bottom row). These are the stimuli for which, in our psychophysical experiment, observers reported the lowest perception of forward-depth. Only the big-gap condition had 2 observers report forward-depth reliably, the no-gap and small-gap conditions had no forward-depth perception (see Fig. 8). Therefore, these conditions represent the most challenging case for the second-order mechanism, which according to the model also predict forward depth perception.

For anti-correlated stimuli where there was no-gap, the horizontal condition had the highest number of observers (3) reporting forward-depth, with none in the no-gap circular, and 1 in the no gap vertical. This was the only no-gap condition where the target and reference did not overlap. This suggests that while the absence of the gap may result in reversed-depth perception in ACRDS for first-order mechanisms, the second-order mechanisms rely on the presence of the gap to reduce the general ambiguity of the stimulus in order to provide the perception of forward-depth. In the case of the anti-correlated circular no-gap condition the complexity of the shape (target surrounded by reference) meant the second-order mechanism was unable to segment the target from the surround. The results of our computational modelling show that the disparity-dependent responses of the second-order channel were sufficient for this to be achieved for both CRDS and ACRDS, and would thus be able to support the perception of depth in the forward direction in both cases. However, a reliable forward-depth signal from this channel did not appear to be available in the ACRDS circle case. The most likely explanation for this is therefore not that the information was not available from the second-order filters, but that the visual system was not able to harness this for the perception of depth-defined shape. It is well-established that the reliable perception of depth depends on relative disparity signals<sup>43–47</sup>, and that extrastriate mechanisms exist which pool disparity signals across space in order for this to be achieved<sup>57–65</sup>. Ziegler *et al.*'s<sup>42</sup> psychophysical results suggest that the spatial pooling of second-order signals in order to resolve disparity defined shape is severely limited. It is this limit that is most likely account for the inability of observers to segment the anti-correlated circular target from its surround in our experiments.

Another possible explanation for our results may lie in the depth cue combination model proposed by Landy *et al.*<sup>66</sup>. Depth cues will vary throughout a scene, and while each individual cue contributes to the estimated weighted average of depth, these will vary in the quality of information they contain<sup>66</sup>. Where a cue does not provide a depth estimate, provides an unreliable estimate, or one that contradicts the other cues, it will be down-weighted. Landy *et al.* suggest a mechanism of combination that statistically estimates depth from weighted averages which depend on the interactions<sup>66</sup> and correlations<sup>15,67</sup> between responses. While their model describes cues in terms of qualities of the stimulus, such as texture, shading and parallax, we have shown that, even within a cue, conflicts might exist, such as between estimates from first- and second-order mechanisms. Therefore it is likely that the average weighted estimate includes the error checking response of the second-order mechanism before interpreting depth. While it is still unknown whether first- and second-order mechanisms co-exist or transition from one to another<sup>31</sup>, our results suggest that when first- and second-order responses are in agreement, depth is perceived normally. Conversely, when the responses from first- and second-order mechanisms conflict, depth would be expected to be perceived at the stage consistent with the direction indicated by the more reliable channel, with the highest decision weight. In the vertical and horizontal stimuli this was the second-order mechanism. However, the spatial complexity of the circular condition suggests that neither first- nor second-order mechanisms are reliable, and therefore the visual system was unable to unambiguously reconcile depth. This would reflect the inability of the second-order channel to reliably signal depth other than in stimuli with very simple spatial variations in surface layout<sup>42</sup>. The results of our modelling show that the outputs of second-order filters could in all cases be used to support the perception of forward-depth. The lack of depth perception in ACRDS is consistent with the proposed inability of the visual system to use the available information for the perception of depth-defined shape.

Reversed-depth perception is not routinely found in centrally presented targets<sup>5,19</sup> however, our results are consistent with a recent paper<sup>68</sup> that found reversed perception of ACRDS in the periphery. Zhaoping and Ackermann<sup>68</sup> account for their findings through reduced (or absent) feedback connections in the periphery, resulting in a reduction of verification for stimulus features (see Zhaoping<sup>68</sup> for a detailed discussion of the feedforward-feedback-verify-weight network model). An important observation to make is that the stimuli used in our study were neither strictly centrally nor peripherally presented, but 'mid-way' between the two positions used by Zhaoping and Ackerman. Since second-order channels are also strongest in central fields, and weaker or absent in the periphery, this may also account for these findings. Therefore, it is also likely that the reversed perception in this condition (circular no-gap) for some participants, was a result of the combination of the part-peripherally located presentation and the complexity of the stimulus.

Since spatially separating the reference and the surround has been found to reduce stereoacuity<sup>17,46</sup>, it has been suggested that the spatial gap may be responsible for the lack of robust reversed-depth perception in ACRDS. Consistent with Kamihirata *et al.*<sup>47</sup> we found reversed-depth when there was no-gap, which diminished as the gap size increased. However, neither the horizontal nor vertical conditions showed this trend. This implies that the absence of the gap is not exclusively responsible for promoting the perception of reversed-depth. The relatively complicated shape of the circular target and its annulus when there is no-gap and thus the opportunity for



overlap also plays a role in reversed-depth, through reducing the effectiveness of the second-order channel to signal forward-depth. Cumming *et al.*<sup>12</sup> found that reversed-depth perception from ACRDS reduced as a function of increased dot density. Perception of reversed-depth has been found at low dot densities (between 1% and 11%)<sup>12,18</sup>. At high dot densities it is more likely that there is overlap between the dots in each eye, increasing the chance of an ambiguous signal.

A clear finding arising from ACRDS studies<sup>12,17,19</sup> is that there is considerable variability between individuals' reported perception of depth. Recent findings for robust reverse perception of ACRDS in the periphery<sup>68</sup> suggest that ACRDS still has an important role to play in allowing us to understand how the visual system learns to process depth. Cumming *et al.*<sup>12</sup> attempted to improve depth perception for ACRDS using training with feedback, but were unsuccessful. They used trial by trial feedback with over 10 000 presentations of ACRDS, yet performance did not improve. Recent advances in perceptual learning research indicates that crucial factors in learning are individual confidence and task difficulty<sup>69–71</sup>. Our results show a very clear difference between participants' confidence in their judgements for CRDS and ACRDS. Perceptual learning to improve depth perception in anti-correlated stimuli may benefit from these advancements by focusing on improving participant confidence.

## Conclusion

How the visual system combines monocular sensory information into a binocular percept is an important question in vision research. Based on the binocular energy response from first-order mechanisms, it has been proposed that reversed-depth perception should occur when presenting an ACRDS target with a CRDS surround. However, these predictions have only taken into account the first-order responses, ignoring the squaring non-linearity of the error checking<sup>15</sup>, 'back up' function<sup>31</sup> of the second-order mechanism. Our results and analyses underline the need to take account of the interactions between multiple multiple binocular mechanisms when predicting the perception of depth. The sometimes conflicting signals available from ACRDS, containing disparities which are not consistent with geometrical considerations, mean that such predictions are by no means straightforward. Our results are consistent with an increasing number of studies that show second-order summation occurs as an independent process taking place after first-order processing, in motion<sup>72,73</sup>, perceived contrast<sup>74</sup> and binocular phase<sup>75</sup> and stereopsis<sup>40</sup>. Finally, while depth perception for ACRDS was mostly ambiguous, the positive slopes found in some conditions indicate a trend for perception in the forward direction, rather than reversed-depth. Furthermore, confidence for ACRDS was poor in every condition, and increasing participant confidence may be the key to improving reversed-depth perception.

## References

- Julesz, B. *Foundations of cyclopean perception*. (U. Chicago Press, 1971).
- Cormack, L. K., Stevenson, S. B. & Schor, C. M. Interocular correlation, luminance contrast and cyclopean processing. *Vis. research* **31**, 2195–2207 (1991).
- Fleet, D. J., Wagner, H. & Heeger, D. J. Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vis. research* **36**, 1839–1857 (1996).
- Allenmark, F. & Read, J. Spatial stereoresolution for depth corrugations may be set in primary visual cortex. *BMC Neurosci.* **12**, P263 (2011).
- Doi, T., Tanabe, S. & Fujita, I. Matching and correlation computations in stereoscopic depth perception. *J. Vis.* **11**, 1–1 (2011).
- Doi, T. & Fujita, I. Cross-matching: a modified cross-correlation underlying threshold energy model and match-based depth perception. *Front. computational neuroscience* **8** (2014).
- Banks, M. S., Gepshtein, S. & Landy, M. S. Why is spatial stereoresolution so low? *J. Neurosci.* **24**, 2077–2089 (2004).
- Nienborg, H., Bridge, H., Parker, A. J. & Cumming, B. G. Receptive field size in v1 neurons limits acuity for perceiving disparity modulation. *J. Neurosci.* **24**, 2065–2076 (2004).
- Ohzawa, I., DeAngelis, G. C. & Freeman, R. D. Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Sci* (1990).
- Ohzawa, I., DeAngelis, G. C. & Freeman, R. D. Encoding of binocular disparity by complex cells in the cat's visual cortex. *J. neurophysiology* **77**, 2879–2909 (1997).
- Cumming, B. G. & Parker, A. J. Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nat.* **389**, 280–283 (1997).
- Cumming, B. G., Shapiro, S. E. & Parker, A. J. Disparity detection in anticorrelated stereograms. *Percept.* **27**, 1367–1377 (1998).
- Martinez, L. M. & Alonso, J.-M. Complex receptive fields in primary visual cortex. *The neuroscientist* **9**, 317–331 (2003).
- Tanaka, H. & Ohzawa, I. Neural basis for stereopsis from second-order contrast cues. *J. Neurosci.* **26**, 4370–4382 (2006).
- Hibbard, P. B., Goutcher, R. & Hunter, D. W. Encoding and estimation of first-and second-order binocular disparity in natural images. *Vis. research* **120**, 108–120 (2016).
- Anstis, S. M. & Rogers, B. J. Illusory reversal of visual depth and movement during changes of contrast. *Vis. research* **15**, 957–IN6 (1975).
- Aoki, S. C., Shiozaki, H. M. & Fujita, I. A relative frame of reference underlies reversed depth perception in anticorrelated random-dot stereograms. *J. vision* **17**, 17–17 (2017).
- Cogan, A. I., Lomakin, A. J. & Rossi, A. F. Depth in anticorrelated stereograms: effects of spatial density and interocular delay. *Vis. research* **33**, 1959–1975 (1993).
- Hibbard, P. B., Scott-Brown, K. C., Haigh, E. C. & Adrain, M. Depth perception not found in human observers for static or dynamic anti-correlated random dot stereograms. *PLoS one* **9**, e84087 (2014).
- Kaufman, L. & Pitblado, C. Further observations on the nature of effective binocular disparities. *The Am. journal psychology* **78**, 379–391 (1965).
- Read, J. C. & Eagle, R. A. Reversed stereo depth and motion direction with anti-correlated stimuli. *Vis. research* **40**, 3345–3358 (2000).
- Rogers, B. J. & Anstis, S. M. Reversed depth from positive and negative stereograms. *Percept.* **4**, 193–201 (1975).
- Tanabe, S., Yasuoka, S. & Fujita, I. Disparity-energy signals in perceived stereoscopic depth. *J. vision* **8**, 22–22 (2008).
- Lippert, J. & Wagner, H. A threshold explains modulation of neural responses to opposite-contrast stereograms. *Neuroreport* **12**, 3205–3208 (2001).
- Nieder, A. & Wagner, H. Hierarchical processing of horizontal disparity information in the visual forebrain of behaving owls. *J. Neurosci.* **21**, 4514–4522 (2001).

26. Read, J. C., Parker, A. J. & Cumming, B. G. A simple model accounts for the response of disparity-tuned v1 neurons to anticorrelated images. *Vis. neuroscience* **19**, 735–753 (2002).
27. Henriksen, S., Cumming, B. G. & Read, J. C. A single mechanism can account for human perception of depth in mixed correlation random dot stereograms. *PLoS computational biology* **12**, e1004906 (2016).
28. Tanabe, S., Umeda, K. & Fujita, I. Rejection of false matches for binocular correspondence in macaque visual cortical area v4. *J. Neurosci.* **24**, 8170–8180 (2004).
29. Janssen, P., Vogels, R., Liu, Y. & Orban, G. A. At least at the level of inferior temporal cortex, the stereo correspondence problem is solved. *Neuron* **37**, 693–701 (2003).
30. Krug, K., Cumming, B. G. & Parker, A. J. Comparing perceptual signals of single v5/mt neurons in two binocular depth tasks. *J. neurophysiology* **92**, 1586–1596 (2004).
31. Wilcox, L. M. & Allison, R. S. Coarse-fine dichotomies in human stereopsis. *Vis. research* **49**, 2653–2665 (2009).
32. Edwards, M., Pope, D. R. & Schor, C. M. Orientation tuning of the transient-stereopsis system. *Vis. Res.* **39**, 2717–2727 (1999).
33. Edwards, M., Pope, D. R. & Schor, C. M. First-and second-order processing in transient stereopsis. *Vis. Res.* **40**, 2645–2651 (2000).
34. Hess, R. F. & Wilcox, L. M. Linear and non-linear filtering in stereopsis. *Vis. research* **34**, 2431–2438 (1994).
35. Langley, K., Fleet, D. J. & Hibbard, P. B. Linear and nonlinear transparencies in binocular vision. *Proc. Royal Soc. Lond. B: Biol. Sci.* **265**, 1837–1845 (1998).
36. Langley, K., Fleet, D. J. & Hibbard, P. B. Stereopsis from contrast envelopes. *Vis. research* **39**, 2313–2324 (1999).
37. McKee, S. P., Verghese, P. & Farell, B. What is the depth of a sinusoidal grating? *J. Vis.* **4**, 1–1 (2004).
38. Schor, C. M., Edwards, M. & Pope, D. R. Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vis. research* **38**, 3057–3068 (1998).
39. Wilcox, L. M. & Hess, R. F. D max for stereopsis depends on size, not spatial frequency content. *Vis. research* **35**, 1061–1069 (1995).
40. Wilcox, L. M. & Hess, R. F. Is the site of non-linear filtering in stereopsis before or after binocular combination? *Vis. Res.* **36**, 391–399 (1996).
41. Wilcox, L. M. & Hess, R. F. Scale selection for second-order (non-linear) stereopsis. *Vis. Res.* **37**, 2981–2992 (1997).
42. Ziegler, L. R. & Hess, R. F. Stereoscopic depth but not shape perception from second-order stimuli. *Vis. Res.* **39**, 1491–1507 (1999).
43. Erkelens, C. & Collewijn, H. Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. *Vis. research* **25**, 1689–1700 (1985).
44. Erkelens, C. & Collewijn, H. Motion perception during dichoptic viewing of moving random-dot stereograms. *Vis. research* **25**, 583–588 (1985).
45. Regan, D., Erkelens, C. J. & Collewijn, H. Necessary conditions for the perception of motion in depth. *Investig. Ophthalmol. & Vis. Sci.* **27**, 584–597 (1986).
46. Cottureau, B. R., McKee, S. P. & Norcia, A. M. Bridging the gap: Global disparity processing in the human visual cortex. *J. Neurophysiol.* **107**, 2421–2429 (2012).
47. Kamihirata, H., Oga, T., Aoki, S. & Fujita, I. A gap between adjacent surfaces deteriorates depth perception based on binocular correlation computation. *J. Physiol. Sci.* **17**, s155 (2015).
48. Cook, M. & Gillam, B. Depth of monocular elements in a binocular scene: The conditions for da vinci stereopsis. *J. Exp. Psychol. Hum. Percept. Perform.* **30**, 92–103 (2004).
49. Gillam, B. & Nakayama, K. Quantitative depth for a phantom surface can be based on cyclopean occlusion cues alone. *Vis. Res.* **39**, 109–112 (1999).
50. Nakayama, K. & Shimojo, S. Da vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vis. research* **30**, 1811–1825 (1990).
51. Tsirlin, I., Wilcox, L. M. & Allison, R. S. Monocular occlusions determine the perceived shape and depth of occluding surfaces. *J. Vis.* **10**, 11–11 (2010).
52. Tsirlin, I., Wilcox, L. M. & Allison, R. S. A computational theory of da vinci stereopsis. *J. Vis.* **14**, 5–5 (2014).
53. Brainard, D. H. The psychophysics toolbox. *Spatial vision* **10**, 433–436 (1997).
54. Kleiner, M. *et al.* What's new in psychtoolbox-3. *Percept.* **36**, 1 (2007).
55. Pelli, D. G. The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision* **10**, 437–442 (1997).
56. McColl, S. L., Ziegler, L. & Hess, R. F. Stereodeficient subjects demonstrate non-linear stereopsis. *Vis. Res.* **40**, 1167–1177 (2000).
57. Thomas, O. M., Cumming, B. G. & Parker, A. J. A specialization for relative disparity in v2. *Nat. neuroscience* **5**, 472 (2002).
58. Bredfeldt, C. E. & Cumming, B. G. A simple account of cyclopean edge responses in macaque v2. *J. Neurosci.* **26**, 7581–7596 (2006).
59. Bredfeldt, C., Read, J. & Cumming, B. A quantitative explanation of responses to disparity-defined edges in macaque v2. *J. neurophysiology* **101**, 701–713 (2009).
60. Zhaoping, L. Pre-attentive segmentation and correspondence in stereo. *Philos. Transactions Royal Soc. Lond. B: Biol. Sci.* **357**, 1877–1883 (2002).
61. Zhaoping, L. Border ownership from intracortical interactions in visual area v2. *Neuron* **47**, 143–153 (2005).
62. Assee, A. & Qian, N. Solving da vinci stereopsis with depth-edge-selective v2 cells. *Vis. research* **47**, 2585–2602 (2007).
63. Umeda, K., Tanabe, S. & Fujita, I. Representation of stereoscopic depth based on relative disparity in macaque area v4. *J. neurophysiology* **98**, 241–252 (2007).
64. Anzai, A., Chowdhury, S. A. & DeAngelis, G. C. Coding of stereoscopic depth information in visual areas v3 and v3a. *J. Neurosci.* **31**, 10270–10282 (2011).
65. Orban, G. A., Janssen, P. & Vogels, R. Extracting 3d structure from disparity. *Trends neurosciences* **29**, 466–473 (2006).
66. Landy, M. S., Maloney, L. T., Johnston, E. B. & Young, M. Measurement and modeling of depth cue combination: In defense of weak fusion. *Vis. research* **35**, 389–412 (1995).
67. Oruç, I., Maloney, L. T. & Landy, M. S. Weighted linear cue combination with possibly correlated error. *Vis. research* **43**, 2451–2468 (2003).
68. Zhaoping, L. & Ackermann, J. Reversed depth in anti-correlated random dot stereograms and central-peripheral difference in visual inference. *bioRxiv* 225532 (2017).
69. Jeter, P. E., Doshier, B. A., Petrov, A. & Lu, Z.-L. Task precision at transfer determines specificity of perceptual learning. *J. vision* **9**, 1–1 (2009).
70. Talluri, B. C., Hung, S.-C., Seitz, A. R. & Seriès, P. Confidence-based integrated reweighting model of task-difficulty explains location-based specificity in perceptual learning. *J. vision* **15**, 17–17 (2015).
71. Lin, Z., Doshier, B. A. & Lu, Z.-L. Mixture of easy trials enables transient and sustained perceptual improvements through priming and perceptual learning. *Sci. reports* **7**, 7421 (2017).
72. Edwards, M. & Nishida, S. Contrast-reversing global-motion stimuli reveal local interactions between first-and second-order motion signals. *Vis. research* **44**, 1941–1950 (2004).
73. Edwards, M. & Metcalfe, O. Independence in the processing of first-and second-order motion signals at the local-motionpooling level. *Vis. research* **50**, 261–270 (2010).
74. Baker, D. H., Wallis, S. A., Georgeson, M. A. & Meese, T. S. The effect of interocular phase difference on perceived contrast. *PLoS One* **7**, e34696 (2012).
75. Zhou, J., Liu, R., Zhou, Y. & Hess, R. F. Binocular combination of second-order stimuli. *PLoS one* **9**, e84632 (2014).

### Author Contributions

J.A. and P.H. conceived the experiment; J.A. conducted the experiment; J.A. and P.H. analysed the data; P.H. and J.A. derived the models; J.A. and P.H. wrote and reviewed the manuscript.

### Additional Information

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018