Owls see in stereo much like humans do

Robert F. van der Willigen

Department of Biophysics, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

While 3D experiences through binocular disparity sensitivity have acquired special status in the understanding of human stereo vision, much remains to be learned about how binocularity is put to use in animals. The owl provides an exceptional model to study stereo vision as it displays one of the highest degrees of binocular specialization throughout the animal kingdom. In a series of six behavioral experiments, equivalent to hallmark human psychophysical studies, I compiled an extensive body of stereo performance data from two trained owls. Computer-generated, binocular random-dot patterns were used to ensure pure stereo performance measurements. In all cases, I found that owls perform much like humans do, viz.: (1) disparity alone can evoke figure–ground segmentation; (2) selective use of “relative” rather than “absolute” disparity; (3) hyperacute sensitivity; (4) disparity processing allows for the avoidance of monocular feature detection prior to object recognition; (5) large binocular disparities are not tolerated; (6) disparity guides the perceptual organization of 2D shape. The robustness and very nature of these binocular disparity-based perceptual phenomena bear out that owls, like humans, exploit the third dimension to facilitate early figure–ground segmentation of tangible objects.

Keywords: binocular vision, discrimination, perceptual organization, spatial vision, shape and contour, visual acuity


Introduction

Stereopsis refers to the compelling, binocularly fused sensation of three-dimensional (3D) shape familiar to most humans (Parker, 2007; Schor, 2003). This is a remarkable feat of evolution, because for stereopsis to be possible, the brain has to pair parts visible from the left eye with non-corresponding parts as seen from the right eye’s vantage point that originate from the same points in 3D space (Barlow, Blakemore, & Pettigrew, 1967). In other words, our solid, single vision of 3D space arises from the brain’s ability to combine two slightly different retinal images.

Nevertheless, how does one specifically test for such visual function? First and foremost, for stereopsis to work it requires so-called horizontal, binocular disparities (hereafter disparities; Barlow et al., 1967). The perceptual independence of disparity-induced 3D shape (Figure 1Ai) can be easily experienced by dichoptic viewing of computer-generated random-dot stereograms (RDSs) for which no familiar—monocular detectable—structure is present (upper diagrams, Figure 1Aii). In practice, the left and right eyes’ images of a single RDS consist of identical black-and-white random-dot patterns. Binocular disparities can then be introduced by a horizontal shift of a square of dots in the center of one eye’s image coinciding with an identical shift, but in opposite direction of corresponding dots in the other eye’s image (lower diagrams, Figure 1Aii). In human observers, able to combine such horizontally disparate images, a floating frontoparallel surface will appear vividly at a different depth plane in an otherwise flat and noisy display of dots (Figure 1Ai). Thus, the computer-based RDS technique, as engineered by Julesz (1960, 1964), provides an invaluable tool to study stereopsis in its purest form.

One understanding about stereopsis is that binocularity evolved predominantly in predators, enabling them to judge the absolute distance of prey (Visual Predation Hypothesis: Cartmill, 1974). In a similar vein, the evolutionary pressure for disparity sensitivity in early primates is assumed to have come from its utility to accommodate accurate depth judgments, enabling them better visual motor control of the forelimbs (Arboreal Hypothesis: Collins, 1921). Although both the visual predation and arboreal hypotheses have overshadowed our thinking about stereopsis for decades (e.g., see Georgieva, Peeters, Kolster, Todd, & Orban, 2009), they are gradually being replaced by a novel scientific rational (for a review, see Nakayama, 2005; Ponce & Born, 2008; Todd, 2004).

An emerging consensus derived from human psychophysics is that disparity-based perceptual phenomena not obviously related to perceived depth are underappreciated (Collett, 1996; Glennerster, Tcheang, Gilson, Fitzgibbon, & Parker, 2006; Heesy, 2009; Nakayama, 2005; Ponce & Born, 2008; Todd, 2004; Tyler, 2004a, 2004b). Particularly, the RDS-based discovery in humans, that without familiarity cues disparity-induced depth perception is still possible, prompted Julesz (1960) to speculate that stereopsis evolved in our primate ancestors to overcome the natural camouflage of insects hiding motionless in the foliage.

Pettigrew (1986) put this idea succinctly when he wrote: “... evolution of stereopsis in the interest of edge...”
detection per se, rather than accurate distance judgments, is the most parsimonious explanation.” He came to this understanding because only birds that need to detect concealed prey (i.e., invisible to monocular scrutiny) have the neural machinery to encode disparities. Ultimately, Nakayama, He, and Shimojo (1995) pushed forward the idea that stereopsis is highly effective in parsing objects into wholes as to determine what in the retinal image actually constitutes the parts and boundaries of a single object. This perceptual phenomenon, called figure–ground segmentation, refers to early visual processing that precedes recognition. It is to this challenging problem of vision that Qiu and von der Heydt (2005) directed their single-cell recordings in macaques. They discovered how

Figure 1. Random-dot stereograms, their construction, and viewing conditions. (A) Schematic diagrams of the (Ai) tested and (Aiii) trained 3D shape categories: “hole” versus “step”, as specified by static and changing disparities, respectively. (Ai) The pictorial outlines and shape transformations of the center and surround are used here to create a sense of depth (“hole” vs. “step”) to the reader but were not present in the stimuli proper. (Aii) Positive disparity-defined Julesz random-dot stereogram (top row) and construction (bottom row). Cross-fusion of the top row half-images reveals a “hole”. Gaps that remain due to the rectangle’s horizontal shift are filled with freshly generated random-dot texture (curved arrows). See main text for details. (Aiii, top row) Pictorial outlines of the motion-in-depth stimulus: a centrally placed stereoscopic target is simultaneously drawn against the outline of a surrounding ground to give the impression of motion in depth. The moving target never traversed the depth plane of the non-disparate surround, and its 3D trajectory deviated 45° from the ground’s midsagittal plane of the stimulus panel. The random-dot texture of the surfaces is omitted for illustrative purposes only. Depending on the eye wherein the dots of the central target region were moved coherently from side to side, the owls were presented with either a “hole” that decreased and increased over time (gray arrows) or, alternatively, a “step” that loomed backward and forward (black arrows). (Aiii, bottom row) “Hole” and “object” stereogram configurations of the motion-in-depth stimulus. Notice that dot movement only occurred in one eye and at the right side of the dashed line. Arrows indicate the initial direction of dot movement at trial onset. (B) Owls, standing on a perch, were trained to view a monitor. A push with their beak on either of the two response bars (“back” versus “front”) could be reinforced with a reward made available from a food dispenser. To ensure that each eye was exposed to the appropriate half-image (dichoptic viewing), the RDSs were viewed through a set of polarized or red/green filters along with a polarizing display. Head movements were monitored in real time through a wired head tracking sensor plus a DC magnetic field generator (gray box).
the neural processing of disparity could enhance the early encoding of border ownership.

If, indeed, stereoscopic encoding of 2D object boundaries bestows a unique perceptual advantage, then this should be evident in the RDS-based stereo performances of any visual system capable of disparity processing. As such, the owl visual system is of great value because it has forward facing eyes combined with binocular pathways and neural machinery, permitting the processing of RDS-based stereo information (Pettigrew & Konishi, 1976b). Although owls can be trained to report 3D shape perception from RDSs (van der Willigen, Harmening, Vossen, & Wagner, 2010), it is not known to what extent their stereo performances compare to that of humans (Harmening, Gobbel, & Wagner, 2007; Harmening, Nikolay, Orlowski, & Wagner, 2009; Iwaniuk & Wylie, 2006; Martin, 2009; Nieder & Wagner, 2001; Pinto & Baron, 2009, 2010). This leaves open the challenge of drawing together RDS-based stereo performance data from both owls and humans.

Here, I show that the owl’s ability to discriminate RDSs is paralleled in many ways by that of man, despite differences in their optical and neural constraints. The robustness and very nature of these perceptual commonalities bears out the idea that the advantage of seeing in stereo is not the encoding of 3D shape per se but rather the early encoding of 2D object boundaries prior to shape recognition.

General methods

Participants

Ethics statement

My tests were purely behavioral and involved no distress or discomfort. Experiments were conducted in accordance with Institutional (RWTH, Aachen) guidelines and the directive of the European Communities Council (86/609/EEC). Owls were cared for under a permit from the Regierungspräsidium Köln (Germany). Human psychophysics on two healthy volunteers was performed after they were being informed about the behavioral procedure. All procedures conformed to the principles and standards expressed in the Helsinki declaration.

Experimental animals

Detailed descriptions of the experimental setup, surgical procedures, and animal care have been given elsewhere (van der Willigen, Frost, & Wagner, 2002, 2003). Briefly, two adult barn owls (Tyto alba) were housed in spacious enclosures in groups of two and kept in excellent health. These birds were tamed by hand rearing and carried a head post that had been fixed to the skull under anesthesia (van der Willigen et al., 2010).

Human subjects

The apparatus, stimuli, and procedures used to test two young adult humans were identical to those used to test the owls, unless specified otherwise. Both emmetropic subjects had 20/20 visual acuity or better and normal stereoscopic acuity (TNO Stereo Test, Laméris Ootech). They participated voluntarily and were highly experienced psychophysical observers but naive as to the study’s purpose. A chin rest was used to minimize head movements. By using monocular Nonius lines, I made sure that the angle of eye convergence was appropriate under the experimental conditions described here. Subjects could take as much time as they needed to judge the stimuli.

Apparatus

Stimulus-presenting apparatus and software

Stimulus sequences, real-time presentation, reinforcing contingencies, and online data processing were controlled by an Octane Silicon Graphics (SGI) workstation. A custom written software application (GLUT OpenGL utility kit; Varsity ANSI_C Developer Package Update 5.3 under IRIX 6.5) was used. A 20-inch (pixel pitch: 0.256 mm), P22-phosphor cathode ray tube (CRT) functioned as the stimulus-presenting panel (model: Sony GDM 20E21). Gamma correction produced a linear relationship between stimulus luminance and the gray level specified by the workstation. Spatial calibration involved the creation of a look-up table that converted the desired visual directions into CRT screen coordinates (van der Willigen et al., 2002).

Real-time head tracking

For online measurements of head movements, a tracking device was used. This tracking system measured six degrees of freedom (position along with orientation) by means of DC magnetic induction. It consisted of a magnetic field transmitter and a small inductive wired sensor (miniBIRD, Ascension Technology; accuracy: 0.18 mm, max transfer rate: 144 Hz). As such, it was possible to monitor the owls’ head position (i.e., viewing distance) as a function of time in Cartesian coordinates. For this purpose, the birds wore a small, wired head tracking sensor (Figure 1B).

Stimuli

Basic RDS configuration

The polarized (see Luminance section below) RDSs, used to train the owls, were segmented into two distinct depth planes (Figure 1Ai, a non-disparate surrounding ground plus a central stereoscopic target), revealing a central, planar surface floating either in front of or, alternatively, behind a larger but similar textured pattern
(3D “step” vs. “hole”, respectively). The texture contained equal amounts, 50%, of randomly positioned dots. All dots were anti-aliased to provide sub-pixel resolution. Visual angles were calculated for the observer’s actual viewing distance as derived online from the head tracking device (see Real-time head tracking section).

Stereogram generation

To create stereoscopic images, one eye’s image was written to the even scan lines and the other eye’s image to the odd scan lines of the CRT, each with a frame rate of 60 Hz. Both anaglyph and polarized RDSs were used. The latter were produced using a liquid crystal modulator (LCM, Tektronix SGS310) placed directly in front of the CRT. The LCM transmitted and blocked the left and right eyes’ images (hereafter half-images) alternately in synchrony with the CRT frame rate when viewed through a set of differently polarized filters. The anaglyph RDSs presented the half-images with complementary colors red and green.

Head position monitoring

Three-dimensional head position and orientation were captured online. A Cartesian coordinate frame of head position space was chosen with the X-axis lying along the horizontal direction, the Y-axis along the vertical direction, the Z-axis along the sagittal direction, and an origin located at the center of the perch (Figure 1B).

Luminance

All experiments were performed under dark room conditions (<0.01 cd·m⁻²). Luminance was measured (model: LS-100; Minolta) at the viewing distance through either polarized or red/green filters plus the LCM. All RDSs appeared against a completely dark background (<0.001 cd·m⁻²). The luminance levels of the gray and black dots of the polarized RDSs were 0.360 and 0.001 cd·m⁻², respectively. Luminance extinction was below 6% (i.e., luminance captured by the blocked eye). In anaglyph RDSs, the CIE chromaticity coordinate for red was (0.622, 0.339), and that for green was (0.286, 0.600). The transmission spectra of the filters were chosen to correspond to the emission spectra of the CRT. Luminance levels of the red, green, and black dots (measured through the red and green filter, respectively, against a <0.001 cd·m⁻² background) were 0.250, 0.060, and 0.001 cd·m⁻², respectively. Luminance extinction of the colored dots was below 3%.

Disparity calculation

Binocular disparities were created by horizontal shifts of the central target relative to the surrounding ground in one eye’s half-image coinciding with identical shifts but in the opposite direction of the corresponding dots in the other eye’s half-image. Calculation of binocular retinal disparity was based on the owls’ actual fixation distance, as measured online by the head tracking device. In this way, target disparities > 0 arcmin create “hole” RDSs, whereas target disparities < 0 arcmin create “step” RDSs, given that the surrounding ground is kept non-disparate. Viewing distance ranged between 300 and 1800 mm. The interpupillary separation or IPS (mean value ± [95% confidence interval] mm) as measured in owl O1 was 40 [37, 43] mm, and that of O2 was 38 [34, 41] mm. The IPS of the human subjects equaled 65 [63, 67] mm.

Observation stimulus

A monocularly displayed, central gray fixation square (0.18 cd·m⁻²; 2.5° × 2.5°), placed against a completely dark background (<0.001 cd·m⁻²), functioned as the observation stimulus. Hence, the observation stimulus was visible only by one eye (left or right, selected with equal probability).

Training and rewarding

Operant conditioning

Details of the operant training procedure have been given elsewhere (van der Willigen et al., 2003). Briefly, owls standing on a perch were required to associate random-dot patterns, organized as either “holes” or “steps”, with one of two response alternatives: “back” or “front”, respectively (Figure 1B). As soon as the owls’ viewing angle was perpendicular to the frontoparallel surface of a completely dark CRT, the owls were presented with the observation stimulus for a randomly changed time interval (<6 s); 100 to 500 ms prior to stimulus onset, the CRT was made dark again. After 9–12 months of intensive training, both owls performed the task reliably (for criterion, see Data collection section).

Training

Initially, the owls were trained to discriminate motion-in-depth stimuli (Figure 1Aiii). At the same time, the animals learned to avoid head movements and were required to fixate the center of the stimulus aperture from a predefined primary position (deviations <1 cm from the X, Y, Z coordinates were allowed) when being presented with the observation stimulus. Gaze deviations <2° from the CRT’s midsagittal plane were permitted. A training session was defined as the amount of trials wherein 12 reinforcements were obtained.

Reward contingencies

During training, the owls were rewarded only for correct responses (conditional reinforcement). When the
owls were tested, however, rewards were given after every trial (unconditional reinforcement). In this way, the animals do not become frustrated when the task is difficult or impossible to solve. At the same time, the animals do not receive feedback about the correctness of their choices. Prior to the experiment proper, the animals were reminded of their task through conditional reinforcement using suprathreshold stimuli.

Data analysis

Computational software

Data analysis was performed by means of an Intel hardware-based (Core_2 Duo CPU_E8500) version of Matlab R2010b (Mathworks).

Data collection

Data were collected in a successive viewing of a balanced but quasi-random sequence of RDSs containing one sample of two possible depth categories (Figure 1A, “step” vs. “hole”). Unless specified otherwise, 100% correct represents perfect performance and 50% correct represents chance performance. \( P_{\text{bin}} \) represents the two-sided, independent binomial probability calculated from the number of correct and incorrect responses with a 0.5 expectation of being correct by chance alone. Reliable performance was defined as \( P_{\text{bin}}(X = 83\%) < 0.0001 \), calculated from the last 70 training trials.

Tests of transfer

Testing of successful transfer was performed as follows. Both the baseline and transfer performances were taken to be the mean value from the percentage correct responses of 12 distinct sessions. Transfer was considered to be successful only if the null hypothesis of equal ranks could not be rejected. The null hypothesis was rejected when \( t_{22} > 2.82 \) (two-sided, Wilcoxon rank-sum test).

Psychometric functions

By means of maximum likelihood estimation, a two-parameter, \( \mu \) and \( \sigma \), cumulative Gaussian function, \( \Psi \), was fitted to the data (as obtained by the method of constant stimuli), which is defined as

\[
\Psi(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx
\]

with \( -\infty < x < \infty \).

Here, \( x \) is the dependent measure; \( \mu \) and \( \sigma \) are the mean position and the standard deviation, respectively.

In the stereoacuity experiment, \( \Psi(x) \) corresponds to the probability of making a “front” response, reflecting that the observer perceived the RDS to be “step” shaped. The \( \mu \) parameter represents the bias toward either negative (\( \mu > 0 \), “step” RDS) or positive (\( \mu < 0 \), “hole” RDS) disparities. The \( \sigma \) parameter was taken to be a measure for assessing the observer’s ability to perform the stereoacuity task. The log-likelihood ratio, based on 100,000 Monte Carlo simulations, allowed verification of the goodness of fit: two-sided \( \chi^2_{\text{deviance}}(13) < 18, p > 0.0788 \) (Wichmann & Hill, 2001).

In the masking experiment, \( \Psi(x) \) corresponds to the probability of making a correct choice (“front” versus “back”). The performance halfway between the lower and higher asymptotes of \( \Psi(x) \) was taken to be a measure for the minimal period needed to correctly judge a given RDS or stereo duration threshold (Tyler, 1991). The log-likelihood ratio, based on 100,000 Monte Carlo simulations, allowed verification of the goodness of fit: two-sided \( \chi^2_{\text{deviance}}(8) < 10, p > 0.1325 \) (Wichmann & Hill, 2001).

Confidence intervals

Confidence intervals of parameters, thresholds, slopes, and means that I report on throughout this study were estimated using Efron’s (1987) non-parametric, bias-corrected, and accelerated bootstrapping algorithm, unless specified otherwise.

Perceptual salience

Compared to the rather unproblematic applicability of RDSs to humans in demonstrating the perceptual salience of disparity-defined figure–ground segmentation, their use in animals is much more difficult (Bough, 1970; Fox, Lehmkuehle, & Bush, 1977). I approached this problem by creating RDSs with a monocular discernible cue—structure-from-motion (Regan & Beverley, 1979)—in addition to binocular disparity. This was achieved by creating translational motion of the boundaries of the figure (coherently moved dots in the center of the RDS) relative to the ground (static dots in the surround of the RDS) in one eye’s image, while leaving unaltered the corresponding elements in the other eye’s image (“hole” vs. “step” RDSs; Figure 1Aiii). As a result, the translational motion of the figure boundaries was defined by changes in binocular disparity over time (upper row, Figure 1Aiii). This type of motion display is called stereoscopic motion or motion-in-depth (Julesz, 1971; Regan, 2000).

In particular, the central dots with a crossed (“step”) pedestal (black arrows, Figure 1Aiii) only moved in the left side of the half-image of the left eye, whereas central
dots with an uncrossed (“hole”) pedestal (gray arrows, Figure 1Aiii) only moved in the left side of the half-image of the right eye. Notably, the direction of coherent motion of the “hole” RDSs always started from left to right, whereas those of the “step” RDSs always started from right to left.

The rationale behind this motion-in-depth task is given as follows. If owls are more prone to pick up the structure-from-motion cue rather than the disparity cue, then one would expect that if this monocular detectable cue is eliminated the owls have to start all over again to master the discrimination of static RDSs. That is, discriminative performance is then expected to drop at chance level (50% correct).

**Stimuli**

**Changing RDSs**

Motion-in-depth RDSs (Regan & Beverley, 1979) were defined by changes in disparity over time (hereafter changing disparity). These polarized stereograms were segmented into two distinct depth planes, revealing a central, planar surface floating either in front of or, alternatively, behind a larger but similar textured pattern (3D “step” vs. “hole”, respectively). The texture contained equal amounts, 50%, of randomly positioned gray and dark dots. The surrounding ground filled a region of $450 \times 450$ dots. The center contained $150 \times 150$ dots. Each dot subtended 1/2 arcmin.

Global motion was created in one eye alone, left or right, and consisted of a sequence of discrete horizontal displacements of a square of $(150 \times 150)$ dots at the left side of a given half-image (see the RDSs of Figure 1Aiii). The displacement step size was 3 dots per frame ($4\degree$/s) oscillating between 9 and 27 dots. When global motion occurred in the left eye, human subjects observed a “step” that loomed backward and forward (black arrows, Figure 1Aiii). Alternatively, when dot motion occurred in the right eye, human subjects observed a “hole” that decreased and increased over time (gray arrows, Figure 1Aiii).

**Static RDSs**

The static RDSs were identical to the changing RDSs, except that the stereoscopic figure region had a fixed position at the center of the surrounding ground.

**Procedure**

In the first condition, two naive owls were trained on a motion-in-depth guided task using changing disparity-defined RDSs (Figure 1Aiii). In the second condition, responses to previously unseen stationary disparity-defined RDSs were compiled to gauge if the owls formed step- and hole-shaped figure–ground categories (Figure 1Ai) based on their previous experience with the motion-in-depth guided task. One must bear in mind that in this novel condition the coherent motion of the center dots is completely absent, making disparity the sole cue to be detected.

**Results/discussion**

The left column of Figure 2 summarizes the acquisition of the trained motion-in-depth task by owls O1 (top panel) and O2 (bottom panel), respectively. Here, 50% correct denotes the performance level as expected to arise by chance alone. Both owls mastered the motion-in-depth task within <39 sessions.

I anticipated that a switch from familiar, changing disparities to novel, stationary disparities necessarily eliminates reliable performance altogether if the owls did not pick up on the disparity information available to them. In this supposed “monocular” scenario, the owls must have
learned to identify that if motion occurred in the left eye, starting from left to right (without appreciating any depth), they should select the “back” response bar. Alternatively, when motion occurred in the right eye, starting from right to left, the owls should select the “front” key.

From Figure 2, it can be seen that the “monocular” scenario did not apply to the owls tested. The abbreviated training sessions (shaded areas, Figure 2) on the static RDSs contrast with the long bout of training sessions on the changing RDSs (white areas, left column, Figure 2). Within 3 sessions (<39 trials) on the static RDSs, both owls attained the pre-change performance level, while having needed 26 to 30 training sessions (>442 trials) on the motion-in-depth guided task. Clearly, the owls’ perceptions from disparity-defined RDSs must have a “step” or “hole” figure–ground quality not unlike that experienced by human subjects.

An important aspect of the rapid transfer from changing disparity to static disparity described here is that it appears as if the owls could not see the RDSs without appreciating their depth configuration. As far as I am aware, this result is without precedent within the comparative literature (Howard, 2002) and is consistent with human stereo vision where horizontal disparity is found to be a particularly salient cue in mediating the perception of relative depth (Julesz, 1960). It raises the intriguing possibility that the way in which humans interpret RDSs is paralleled by how owls perceive such stereoscopic stimuli.

Relative vs. absolute disparity

A trenchant characteristic of stereopsis in primates is its dependence on directional differences of two or more points in space (hereafter relative disparity) rather than the absolute position of a single point in space from two vantage points (hereafter absolute disparity; Westheimer, 1979). The importance of the distinction between absolute and relative disparities is reflected by the finding that stereopsis informs humans (Collett, 1996; Goutcher & Hibbard, 2010; Schor, 2003; Todd & Norman, 2003) and macaque monkeys (Cumming & Parker, 1999; Haeftner & Cumming, 2008; Neri, Bridge, & Heeger, 2004; Prince, Pointon, Cumming, & Parker, 2000) essentially about 3D shape rather than absolute depth.

The distinction between relative and absolute angular measures of stereoscopic vision is not only important in terms of perceived depth (Erkelens & Collewijn, 1985b) but also in terms of changes in eye position (Erkelens & Collewijn, 1985a). For instance, primates habitually use vergence eye movements to align their eyes on a single surface. In particular, humans can rapidly correct misalignments by sensing the difference in position between the two retinal images of a given surface without sensing depth (Masson, Büsettini, & Miles, 1997). In marked contrast, barn owls habitually display slow, small (<2°) and unidirectional changes in vergence position of their eyes (du Lac & Knudsen, 1990; Knudsen, 1989; Pettigrew & Konishi, 1976a). This lack of fast vergence could mean that a given retinal disparity is uniquely related to the physical distance of an object, enabling barn owls to rely on absolute disparities to scale viewing distance.

To draw a clear distinction between sensitivities to relative and absolute disparities, I used two distinct sets of RDSs: (1) center–surround RDSs and (2) wide-field, uniform RDSs (i.e., binocular random-dot patterns without a fixed reference frame).

Stimuli

Relative disparity-defined RDSs

Polarized α RDS pairs (Figure 3A) were segmented into two distinct depth planes (a surrounding ground plus a central stereoscopic target), revealing a central, planar surface floating either in front (~16 arcmin) or, alternatively, behind (+16 arcmin) a larger but similar textured non-disparate pattern (3D “step” vs. “hole”, respectively). The texture contained equal amounts, 50%, of randomly positioned gray and dark dots. The surrounding ground filled a region of 450 × 450 dots. The center contained 150 × 150 dots. Each dot subtended 1 × 1 arcmin.

Two novel, transfer RDS pairs, β and γ, were constructed by changing both sign and magnitude of the center disparity in the α baseline RDS pair (Figure 3A), while leaving the relative disparity between the center and surround intact. The β RDS pair was constructed by placing the “step” behind (center 16 arcmin/ground 32 arcmin) and the “hole” in front (center −16 arcmin/ground −32 arcmin) of the depth plane of the stimulus-presenting panel. The γ RDS pair was constructed by placing the ground of both the “step” (center 0 arcmin/ground 16 arcmin) and “hole” (center 0 arcmin/ground −16 arcmin) RDSs at the same depth plane of the stimulus-presenting panel.

Absolute disparity-defined, ground-only RDSs

As to form wide-field stimuli of 30° × 30°—basically similar to the ones used by Erkelens and Collewijn (1985b)—the viewing distance was reduced to 30 cm. The half-images of these polarized RDSs contained 100 × 100 dots (dot size: 18 × 18 arcmin; 50% dot density), which were presented on a flat CRT screen, 70° × 54° in size. As to avoid any relative disparity information, center and surround were non-disparate, thus creating “ground-only” RDSs. Anti-aliasing allowed dot displacement at arbitrary positions between integral pixel locations by placing the center of maximum luminance at the desired
location using a Gaussian function. That is, a dot consisted of a single blob in the shape of 2D Gaussian luminance profile. The individual blobs had a mean luminance of 0.360 cd m^{-2} and a standard deviation of 4 arcmin. The dots were displayed in red because the LCM works best at long wavelengths (luminance extinction <2%). At stimulus onset, disparity was set to zero and the dots in both half-images were moved sideways with equal velocities (varying from 0.5 up to 2°/s in steps of 0.5°/s; selected with equal probability) but in opposite directions.

Figure 3. Evidence of relative disparity-based stereopsis in owls O1 and O2. (A) Systematic change of absolute depth in the trained baseline RDSs to test for relative disparity-based discriminative performance. Drawn are three top-view depth diagrams of the trained baseline, \( \alpha \), and novel transfer, \( \beta + \gamma \), RDS stimulus pairs ("step" vs. "hole"). The black and white rectangles represent the central, target region of "step" (left column) and "hole" (right column) shaped RDS, respectively. Gray rectangles represent the ground region of the respective RDSs. The dashed lines coincide with the absolute distance of the CRT screen. Notice that within each depth category ("step" vs. "hole") the relative center–surround disparity remains constant. (B) The bars show the mean percentage \( \pm \) SD of correct performance from 2 \( \times \) 12 sessions obtained with the trained baseline RDS pair \( \alpha \) (shaded bars, O1: 92\% \pm 6%; O2: 91\% \pm 5%) and the two novel, transfer RDS pairs \( \beta + \gamma \) (open bars, O1: 87\% \pm 7%; O2: 88\% \pm 7%), respectively. The error bars indicate SD; \( N \) denotes number of trials. The gray dashed lines at 50\% correct mark the performance level expected when the owl would be guessing; n.s. signifies that performance with the novel RDSs (open bars) does not significantly deviate from baseline levels (shaded bars). Notably, immediate transitive inference (i.e., without learning) to the transfer depth profiles suggests that the owls responded selectively to relative depth rather than absolute depth. (C) Histograms (mean percentage \( \pm \) SD) represent the same data as depicted by the open bars in (B) but now for the \( \beta \) (O1: "step" 89\% \pm 7, "hole" 85\% \pm 11; O2: "step" 89\% \pm 8, "hole" 89\% \pm 9) and \( \gamma \) (O1: "step" 82\% \pm 11, "hole" 86\% \pm 11; O2: "step" 87\% \pm 13, "hole" 87\% \pm 8) RDS trials separately. (D) The bars (pointing upward) show the mean percentage \( \pm \) SD of correct performance from 12 sessions (each with 12 trials) for three distinct viewing conditions with the ground-only RDSs: (i) "motion" condition; (ii) "absolute" condition; (iii) "relative" condition. The trials of each condition were randomly interleaved (see main text for details). Notably, both owls were only unresponsive when exposed to the ground-only RDSs ("absolute" condition), as can be seen from the high percentage of trials (bars pointing downward) in which the owls were unresponsive.
Procedure

Relative depth test

Sensitivity to the center–surround depth order of the RDSs was tested according to a blocked discrimination-transfer procedure (van der Willigen, Frost, & Wagner, 1998). A daily experiment consisted of a sequence of 4 sessions (of 12 trials each: 6 “steps” plus 6 “holes”). In the first two sessions, the baseline stimuli, \( a \), appeared and were replaced by the novel transfer stimuli, \( \beta + \gamma \), in the third and fourth sessions, each presented with equal probability. The surrounds of the \( a \) RDS pair were non-disparate, whereas center disparities were either +16 or −16 arcmin. The two transfer RDS pairs were constructed from the \( a \) RDS pair. In \( \beta \), center disparities were changed to −16 and +16 arcmin, respectively, while surround disparities were changed to −32 and +32 arcmin, respectively. In \( \gamma \), the central target regions were made non-disparate, whereas surrounding ground disparities were changed to −16 and +16 arcmin, respectively.

Absolute depth test

Ground-only RDSs were used to determine if a change in absolute disparity across the visual field is an adequate cue for the perception of a change in absolute depth. During testing, three stimulus conditions were randomly intermixed: “absolute” condition, “relative” condition, and the “motion” condition. In the “absolute” condition, the backward and forward moving “ground-only” RDSs were presented with equal probability. In the “relative” condition, relative depth in the “ground-only” stimuli was restored by the presence of a non-disparate, locally defined (2° × 6° in size; 100% white dots, 0.36 cd m\(^{-2}\)). In the “motion” condition, the global motion of the random dots became noticeable because in this case only one of the half-images of the “ground-only” RDSs was made visible. Each condition was presented \( N = 144 \) times with equal probability (3 × 12 sessions of 12 trials each).

Results/discussion

Testing for relative disparity-based stereopsis

The systematic manipulations in absolute disparity of the center in the baseline \( a \) RDS pair (Figure 3A), to, respectively, the opposite sign, \( \beta \), and zero disparity, \( \gamma \), while leaving unchanged the disparity difference between center and surround make clear predictions about the type of angular measure that is used. Performance levels should decrease for the \( \beta \) condition if discrimination is dependent on the absolute disparity of the center. That is, 100% becomes 0% when “steps” are mistaken for “holes”. Performance reversal should occur also for the \( \gamma \) condition if discrimination is dependent on the RDS’s mean disparity, but performance should drop to chance level if the owls relied on the center disparity alone. Contrastingly, no performance change is expected when the owls’ generalizations of the \( \beta \) or \( \gamma \) conditions were driven by the center–surround depth order (i.e., relative disparity).

When confronted with the novel RDSs (Figure 3A, \( \beta + \gamma \)), both owls exhibited high performance levels significantly deviating from random responding (Figure 3B, open bars, O1: \( N = 144 \) trials, \( P_{\text{bin}}(X = 87\%) < 0.0001 \); O2: \( N = 144 \) trials, \( P_{\text{bin}}(X = 88\%) < 0.0001 \)). Importantly, Wilcoxon rank-sum testing yielded high values (O1: \( t_{22} = 1.5, p > 0.15 \); O2: \( t_{22} = 1.5, p > 0.18 \)) signifying that the performance with the novel RDSs (Figure 3B, open bars) does not deviate from baseline performance (Figure 3B, shaded bars). Finally, one-way analysis of variance reveals that the performance levels between the “step” and “hole” depth configurations across the \( \beta \) and \( \gamma \) trials (O1: \( F_{(3,44)} = 0.14, p = 0.94 \); O2: \( F_{(3,44)} = 0.16, p = 0.92 \)) were indistinguishable (Figure 3C). Evidently, the owls responded selectively to the difference in absolute disparity between the center and its surrounding ground of the respective RDSs.

Testing for absolute disparity-based stereopsis

By deployment of wide-field RDSs, Erkelens and Collewijn (1985b) showed that absolute disparity changes—applied evenly over the visual field—can evoke eye vergence movements but do not produce sensations of changing depth in humans. Here, I used wide-field, ground-only RDSs to determine if owls are sensitive to changes in absolute disparity. That is, the dots in both half-images were moved sideways with equal velocities but in opposite directions. Thus, absolute but not relative disparity will change over time. Notice that this task is rather similar to the owl’s original task, as described in Figure 1B, but now a frame of reference is lacking. In this way, the owls were forced to indicate whether the presented surface moved closer or grew more distant over time relative to the point of view rather than relative to a surrounding frame of reference.

When the filters of the spectacles were of the same polarization, inducing monocular occlusion (for a more detailed explanation, see the Controls section below), both owls immediately responded (<700 ms) to the onset of the coherent motion in the ground-only RDSs (Figure 3D, “motion” condition). This monocular-based performance did not deviate significantly from guessing (O1: \( N = 144 \) trials, \( P_{\text{bin}}(X = 57\%) > 0.01 \); O2: \( N = 144 \) trials, \( P_{\text{bin}}(X = 44\%) > 0.008 \)). Thus, although the owls were not trained to respond to lateral motion, they responded indiscriminately to the presence of lateral motion and could not use this cue to “false” distinguish between the ground-only RDSs in terms of a surface that moved closer or grew more distant over time.

In marked contrast, the owls only rarely responded after the motion onset of the ground-only RDSs (2 times out of \( N = 288 \) trials) when viewed stereoscopically (Figure 3D, “absolute” condition). Thus, although the owls were given up to 10 s to respond, the changes in absolute disparity
went unnoticed. Finally, when a small \((2^\circ \times 6^\circ)\) non-disparate but locally defined bar (100% white dots, 0.36 cd m\(^{-2}\)) was introduced into the owls’ line of sight while being presented with a ground-only RDS (Figure 3D, “relative” condition), high performance levels were reestablished in both owls \((O1: N = 144\) trials, \(P_{\text{bin}}(X = 87\%) < 0.0001; O2: N = 144\) trials, \(P_{\text{bin}}(X = 86\%) < 0.0001)\). Clearly, by the reinstatement of relative disparity, a sense of motion-in-depth was restored in the owls. In conclusion, global changes in absolute disparity fail to give owls, like humans (Regan, Erkelens, & Collewijn, 1986; Regan, 2000), a sensation of motion-in-depth, making disparity-based depth scaling improbable.

The absence of a depth sensation from full-field random-dot stereograms, as used here, may seem insignificant because these virtual stimuli do not expand in the manner that a tangible object would when it moves closer to an observer. That is, a real textured surface, moving in depth, would generate a changing pattern of both horizontal and vertical disparities along with changes in looming, such as dot size, density, and luminosity. In the absence of looming, the brain may interpret absolute disparity-induced vergence as a lapse in eye movement control, which leaves it no other alternative as to suppress this erroneous “depth cue” (Brenner & Smeets, 2000; Brenner, Van Den Berg, & Van Damme, 1996). Although this confounding effect of vergence may play a role in humans (Howard, 2008), it is likely to be of a lesser problem in the owl as it lacks fast vergence movements (du Lac & Knudsen, 1990; Knudsen, 1989; Pettigrew & Konishi, 1976a). Recently, Wismeijer and Erkelens (2009) suggested that looming enhances our sensation of motion-in-depth only because it strengthens vergence and not some other cue that could be attributed to a source directly involved in the motion itself. This previously unrecognized role of looming in human vision makes it doubtful that the absence of it abolishes the sensation of absolute disparity-induced depth.

### Acute stereopsis: Owl vs. man

Stereopsis in primates is exquisitely sensitive (Tyler, 1991). For humans and macaque monkeys, RDS-based stereacuity can be on the order of 2–3 arcsec (Stevenson, Cormack, & Schor, 1989) and 10–20 arcsec (Prince et al., 2000), respectively, but only for foveally presented stimuli of maximal contrast. Still, the minimum angular resolution (MAR) of the primate eye approximates 60 arcsec, equaling 2 times the diameter of a single foveal photoreceptor (Wassle & Boycott, 1991). Stereacuity in primate is, therefore, a clear example of hyperacuity: the ability to judge the relative position of a visual feature to within a fraction of the eye’s resolving power (Westheimer, 1981).

### Stimuli

Polarized RDSs were segmented into two distinct depth planes (Figure 1Ai, a surrounding ground plus a central stereoscopic target), revealing a central, planar surface floating either in front of or, alternatively, behind a larger but similar textured pattern (3D “step” vs. “hole”, respectively). The texture contained equal amounts, 50\%, of randomly positioned gray and dark dots. Stereogram presentation occurred in either a dynamic or static fashion.

#### Static RDSs

In static RDSs, the surrounding ground filled a region of \(450 \times 450\) dots. The center contained \(150 \times 150\) dots. Each dot subtended \(1 \times 1\) arcmin.

#### Dynamic RDSs

Owing to technical limitation, dynamic RDS dots subtended \(1 \times 2\) arcmin (center: \(150 \times 75\) dots; surround: \(450 \times 225\) dots). Dot positions in static RDSs were updated from trial to trial, whereas dot positions in dynamic RDSs were updated in synchrony with the refresh rate of the monitor. In dynamic RDSs, each eye’s image was replaced every 16.67 ms.

### Procedure

#### Stereacuity (method of constant stimuli)

Threshold performance was determined using a forced choice back–front task in combination with a randomized block design (van der Willigen et al., 1998). Accordingly, stimuli were presented in separate blocks, containing a counterbalanced but quasi-random sequence of “holes” and “steps” spanning a predetermined range of stimulus levels. Both static and dynamic polarized RDSs of maximal contrast were used to determine stereacuity. The RDS’s center disparity was chosen from 15 possible values corresponding to \(-7, -6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6,\) or \(7\) arcmin. For the parameterization of the fitted performance function, see the Psychometric functions section.

### Results/discussion

The owls’ acute stereoscopic capabilities were assessed from the psychometric functions shown in Figure 4A. Here, the probability of “back” responses (i.e., the RDSs was perceived to be a “hole”) is plotted as a function of the center–surround disparity magnitude, using negative values to indicate crossed disparities (“step” RDSs) and positive values for uncrossed disparities (“hole” RDSs).
Stereoacuity to static RDSs

I found two hallmarks of hyperacute stereopsis, which are highly compatible with both macaque (Prince et al., 2000) and human stereo performances (Howard & Rogers, 2002). First, discriminative behavior was directly coupled with the signs and magnitudes of the employed relative disparities (Figure 4Ai). Performance ranged from nearly zero for the largest negative disparities to almost 100% correct for the largest positive disparities. Second, the owl’s best stereoacuity $V_A$, 1.8 arcmin, implies that each eye must have resolved a visual angle of 0.9 arcmin (see $\frac{\Delta \gamma}{2}$ in Figure 4B). However, application of the Nyquist–Shannon sampling theorem on the barn owl’s retinal ganglion cell density predicts a MAR on the order of 4 arcmin (Bravo & Pettigrew, 1981; Wathey & Pettigrew, 1989). This value is in close agreement with that derived from the barn owls’ pattern electroretinogram (Ghim & Hodos, 2006) but deviates by a factor of 2.5 when the MAR is estimated behaviorally ($\approx$10 arcmin; Harmening et al., 2007, 2009). In other words, the measured stereoacuity values exceed the eye’s predicted anatomical resolving power by at least a factor MAR/$\Delta \varphi = 4.4$. In addition, notice that the owls were under stimulus control, since both animals showed negligible response biases: O1: $-9 [-22, +23]$ arcsec versus O2: $-3 [-22, +28]$ arcsec ($\mu \pm [95\% \text{ confidence interval}]$ arcsec). In other words, the

Figure 4. Hyperacute stereopsis under symmetric viewing in owls O1 and O2. (Ai) Stereoacuity determined with static RDSs. Data points, of 50 trials each, represent the percentage of trials wherein the stereoscopic target was judged as lying behind a surrounding, non-disparate ground as a function of retinal disparity. The owls were trained to press the “back” response bar when perceiving a “hole”; otherwise, the “front” response bar should be selected. As such, I measured the owl’s ability to categorize between step- and hole-shaped RDSs (as illustrated in Figure 1Ai), which represent negative and positive target disparities, respectively. The vertical bars mark the Bernoulli 95% confidence intervals. The black psychometric curves are cumulative Gaussian functions (Equation 1) fitted to the data. The SDs of the individual curves ($\sigma$, gray insets) represent a measure of stereoacuity (arcmin; i.e., low values equate to high sensitivity, $1/\sigma$). Notice that response biases, as derived from the mean position of the psychometric curves (dashed vertical lines), are negligible. (Aii) As in (Ai) but now for stereoacuity determined from dynamic RDSs. Importantly, prior to testing the owls were never exposed to dynamic RDSs. (B) Symmetric convergence for two points in space $P_1$ and $P_2$ showing the human geometrical relationship of the vergence angles $\alpha$ and $\beta$ to the visual angle $\Delta \varphi$ and the depth displacement, $\Delta d$. The base of the black and gray triangles represents the interpupillary separation or IPS with the nodal points of the two converged schematic eyes at the corners. $D$ represents the viewing distance to fixation point $P_1$. Relative difference in absolute disparity between $P_1$ and $P_2$ is given by $\alpha - \beta$, or $2\Delta \varphi$. For details about the owls’ binocular viewing geometry (gray inset), see main text.
owls are equally sensitive to both positive and negative disparities.

For the purpose of directly comparing the stereoscopic resolving power between the owl and man, stereoacuity was measured in two adult humans (male, 23 and 25 years of age) highly experienced in viewing polarized RDSs. I determined that stereoacuity \( (\sigma \pm [95\% \text{ confidence interval}]) \) arcpmin from static RDSs was hyperacutely in both observers (H1: 0.28 [0.26, 0.30] arcpmin; H2: 0.27 [0.26, 0.28] arcpmin), as it has been typically measured with high density, static RDSs of maximal contrast (Backus, Fleet, Parker, & Heeger, 2001; Tyler, 1991). A lower limit of 0.26 arcpmin implies that each eye must have resolved a visual angle, \( \Delta \phi \), as small as 0.13 arcm. Nevertheless, the MAR of the human eye is on the order of 1 arcm, as calculated from the peak retinal ganglion density of the human retina, assuming a retinal magnification factor of 0.282 mm/deg (Curcio & Allen, 1990; Thibos, Cheney, & Walsh, 1987). Consequently, the stereoacuity measured here exceeds the anatomical resolving power of the human eye by a factor \( \text{MAR}/\Delta \phi = 7.7 \).

The facilitating effect on the angular resolution of human stereopsis, as determined here, is maximally 2 times the value as calculated for the owls (MAR/\( \Delta \phi = 4.4 \)). This corresponds well with difference in the binocular viewing geometry (Figure 4B) that exists between the owl (gray inset) and human visual system. Namely, the eyes of the owls were on average 0.58 times closer together than that of the humans tested (see the Disparity calculation section). Given that relative disparity is an angular measure (Figure 4B), it becomes approximately equal to \( \text{IPS} \cdot \Delta d \cdot D^{-2} \) when the viewing distance, \( D \), is large by comparison with physical depth, \( \Delta d \) (Cormack & Fox, 1985). Thus, geometry predicts that 1/0.58 = 1.7 times the physical depth is required to obtain a given relative disparity for owls as for humans. In turn, this hypothetical value corresponds well with my measurements, since the human MAR/\( \Delta \phi \) ratio divided by that obtained for the owl amounts to 7.7/4.4 = 1.8. Clearly, the resolving power of stereopsis in the owl, as observed here, is in close quantitative agreement with that of man when taking into account both the MAR and IPS of the respective visual systems.

To put emphasis on the uniqueness of the owls’ binocular viewing geometry, I have provided a schematic drawing of it (gray inset, Figure 4B) on the right side of the schematic representing the human binocular viewing geometry. The divergent angle (≈62°) between the virtually immobile eyes (du Lac & Knudsen, 1990) in the owls’ skull allow for binocular vision only on the temporal part of the retina in each eye. This is the part where we find the fovea (upward pointing black arrows), which coincides with a retinal area of elevated ganglion cell density (Wathey & Pettigrew, 1989). As a result, the visual axes (VAs) of the owls’ eye run almost parallel, whereas the optical axes (OAs) make an angle of approximately 62 degrees (Oehme, 1961). Notice also that because the eyes of the owl are closer together compared to the human condition, the relative disparity \( \Delta \phi \), as evoked by the depth displacement \( \Delta d \), is substantially smaller.

**Stereaoacuity to dynamic RDSs**

The owls’ immediate (i.e., without prior exposure) and highly reliable discrimination of dynamic RDSs (Figure 4Aii) reveals a fundamental aspect of stereoscopic processing. With dynamic RDSs, each subsequent binocular pair of frames contains dots masking the previous frames just by randomly changing the spatial distribution of the random dots. In view of these theoretical considerations, it was Julesz (1994), the inventor of the computer-generated RDS, who pointed out that at high frequencies (≥60 Hz) masking necessarily prevents top-down processes from penetrating down in time. Thus, stereoscopic processing in the owl is likely to occur prior to object recognition. This finding is not novel for animals, however, as it has been described for non-human primates (Poggio, Motter, Squatrito, & Trotter, 1985).

In both owls, stereoacuity \( (\sigma \pm [95\% \text{ confidence interval}]) \) arcpmin measured with static (Figure 4Ai) RDSs (O1: 1.8 [1.2, 2.6] arcpmin; O2: 2.5 [2.1, 3.1] arcpmin) was on average a factor 0.7 lower when measured from dynamic (Figure 4Aii) RDSs (O1: 3.3 [2.6, 3.6] arcpmin; O2: 3.4 [2.7, 3.7] arcpmin). This may be due to the existence of some temporal incongruence in the disparity signal of the stimulus-presenting system. That is, imperfections in the half-image synchronization will render an elevated estimate of human stereaoacuity (Tyler, 1991).

Nonetheless, under the dynamic condition observers could, in principle, benefit from additional statistical independent samples of the disparity cue. As a control, I therefore tested two human subjects under identical conditions. On average, stereoacuity \( (\sigma \pm [95\% \text{ confidence interval}]) \) arcpmin from dynamic RDSs (H1: 0.54 [0.52, 0.60] arcpmin; H2: 0.44 [0.41, 0.48] arcpmin) was a factor 0.6 lower than was determined with static RDSs (H1: 0.28 [0.26, 0.30] arcpmin; H2: 0.27 [0.26, 0.28] arcpmin). Conceivably, diminished performance from dynamic RDSs compared to the static condition is caused by systematic imperfections in the half-image synchronization of the stimulus-generating apparatus.

**Processing time: Owl vs. man**

The effectiveness of disparity to recognize 3D shape can be characterized by the rate at which stereo performance declines when the exposure time is reduced (Julesz, 1994;
Tyler, 1991). Particularly, by the deployment of ambiguous RDS-based (Figure 5A) backward masking, Julesz (1964) and later Uttal, Fitzgerald, and Eskin (1975) obtained estimates of the minimum unimpeded period, ≈50 ms, that humans require to reliably discriminate 3D shape from RDSs. Notably, this time constant is too short to evoke a vergence eye movement in humans, implicating that ≈50 ms is sufficient for binocular input to evoke sufficient neural activity to ultimately reach the stereoscopic threshold for 3D shape recognition.

For the purpose of directly comparing the time required by the owl for a given disparity magnitude to reach threshold, I replicated Julesz’s (1964) masking experiment.

**Stimuli**

**Masking stimuli**

Ambiguous RDSs (Figure 5A), portraying two depth configurations (“step” versus “hole”) simultaneously (Julesz, 1964), can erase the afterimages of RDSs depicting a single, unambiguous depth configuration: “step” or “hole”. The non-disparate center regions of the noise mask half-images included a grid-like structure of vertical stripes with constant periodicity (comparable to the ambiguous middle rectangle as depicted in the RDS of Figure 5A). The periodicity of vertical stripes between the left and right half-images was shifted by one period. The horizontal

---

Figure 5. Stereoscopic processing time: owl (O1 plus O2) vs. human (H1 plus H2). (A) RDS that, when viewed crossed fused, contains a top rectangle perceived behind the surround (“hole”), a bottom rectangle perceived in front of the surround (“step”), and an ambiguous, central rectangle perceived as either a “step” or “hole” (adapted after Julesz, 1964). The ambiguous nature of the central rectangle becomes apparent when it is viewed after prior exposure to one of the non-ambiguous RDS regions. For instance, if the top rectangle is viewed first, then the ambiguous rectangle is seen receded to the back. Contrastingly, if the bottom rectangle is viewed first, then the ambiguous rectangle is seen in front of the surround. (B) Data points, comprising 2 × 30 trials each, represent the percentage of trials in which an ambiguous stereoscopic target was judged to have the same depth configuration (“step” or “hole”) as the preceding, unambiguous RDS. The backward-masking delays ranged from 10 to 100 ms. Stimulus viewing duration equaled 20 ms (gray lines) or 40 ms (black lines). Data were pooled over two owls (O1 and O2; left side panel) or two humans (H1 and H2; right side panel). The sigmoid-shaped psychometric curves are fitted cumulative Gaussian functions (Equation 1). The thresholds at ≈75% correct of the individual curves (μ, gray insets) mark the minimal stereoscopic processing time needed to discriminate “step” vs. “hole” RDSs. Vertical bars mark the Bernoulli 95% confidence intervals. Gray data points are displaced slightly to right side relative to the black data, for illustrative purposes only. Notably, the 95% confidence intervals of the thresholds, as determined for each observer separately (see main text), do not deviate significantly; as well as to those obtained for the 20 (gray lines) and 40 (black lines) ms stimulus exposure time.
length of this period equaled the shift that existed between the disparate center regions of the “step”/“hole” half-images, constituting the ambiguous central region of the RDSs. The surrounding ground in all (polarized, static) RDSs filled a region of 450 × 450 dots. The center encompassed 150 × 150 dots. Each dot subtended 1 × 1 arcmin. The width of the noise mask equaled 150 pixels plus one period.

Procedure

**Stereo-erasing masking (method of constant stimuli)**

A stereo-erasing or “backward-masking” episode (800 ms in duration) of an ambiguous RDS (noise mask) followed the prior display of suprathreshold “step” or “hole” configured RDSs (target). By varying this target-to-mask time interval (i.e., the elapsed time prior to masking), the unimpeded exposure time for the target is carefully controlled for. The masked target RDSs were presented in blocks containing a counterbalanced but quasi-random sequence of suprathreshold “holes” and “steps” spanning a predetermined range of backward-masking delays (i.e., interstimulus intervals), chosen from 10 possible values ranging from 10 up to 100 ms in steps of 10 ms. The duration of the unambiguous target RDS equaled 20 or 40 ms. Viewing distance was 180 cm. Most notably, to correct for the difference in IPS between humans and owls, suprathreshold center disparity of the “step”/“hole” RDSs was set to −5 or +5 arcmin and −10 or +10 arcmin, respectively. For the parameterization of the fitted performance function, see the Psychometric functions section.

**Results/discussion**

The depth inducing effect—as first described by Julesz (1960)—of a stereoscopic image on ambiguous RDSs can be appreciated from Figure 5A. Note, however, that because of the simultaneous presence of both the unambiguous RDS and the “step”/“hole” stereoscopic images, the perception of the mask is almost exclusively dependent on the observers’ vergence posture. In the experiment proper, however, the actual exposure time of the stereoscopic image was very brief (<50 ms) as to avoid the influence of a “biased” vergence angle.

Steroscopic processing times were assessed from the psychometric functions shown in Figure 5B (owl vs. human). Here, the probability of correctly indicating the depth configuration of “hole” and “object” RDSs (percentage correct) is plotted as a function of the backward-masking delay in ms. Here, the performance halfway between the lower and higher asymptotes of the psychometric function (denoted as μ in the gray insets of Figure 5B) was taken to be a measure for the minimal period needed to correctly judge a given RDS or stereo duration threshold.

I found that depth judgments masked backward with a delay <60 ms are severely degraded. Most notably, the threshold (μ ± [95% confidence interval] ms), as determined in the owls (20-ms exposure: 52 [47, 56] ms; 40-ms exposure: 55 [51, 60] ms), did not deviate significantly from those determined for the two human subjects (20-ms exposure: 52 [45, 59] ms; 40-ms exposure: 52 [47, 58] ms).

Unquestionably, stereopsis in the owl can develop as rapidly as observed in humans and must be the result of early binocular processing of visual information that does not critically depend on convergence eye movements. Most notably, the decisive temporal parameter in both owls and humans is not the exposure time of the stimulus proper (i.e., the unambiguous RDS) but the time given during which unimpeded (unmasked) processing of the disparity signal can occur (backward-masking delay).

As such, the human data of Figure 5B (right panel) replicate the findings of Uttal et al. (1975). Note also that backward masking with ambiguous RDSs has never been attempted before in an animal species (Collett & Harkness, 1982; Howard, 2002) Thus, contrary to mainstream belief (see, e.g., Nadler, Angelaki, & DeAngelis, 2008), the data shown in Figure 5B open the intriguing possibility that the algorithm underlying stereoscopic processing in the owl possesses the same level of sophistication as found in humans and primates.

**Upper disparity limits**

In humans, perceived differences in depth do not increase indefinitely with disparity. In particular, under stabilized conditions of vergence, the fusional limit determined from RDSs, with the dots covering 50% of the image space, may extend up to 112 arcmin for positive (crossed) disparities and 67 arcmin for positive (uncrossed) disparities (Erkelens, 1988). Notice that the specific details about the RDS configuration are highly relevant, since upper depth estimates in humans critically depend on both dot size and density (Glenerster, 1998).

Notably, the condition of stabilized vergence is an important experimental constraint when comparing the upper depth limit of humans with that of the owl. First, briefly exposed RDSs (with disparities up to 2°) are known to evoke fast vergence responses without the impression of depth even after the stimulus has been masked or taken away (Howard & Rogers, 2002; Masson et al., 1997). Second, barn owls cannot produce fast vergence eye movements (du Lac & Knudsen, 1990; Knudsen, 1989; Pettigrew & Konishi, 1976a).

To assess the upper range of relative disparities over which stereoscopic depth perception does occur in the owl, I used static, polarized RDSs with 50% dot density.
Stimuli

Configuration and presenting method

Presentation of the static RDSs occurred in a polarized fashion (see General methods section). RDS configuration was identical as described in Experiment 3: Acute stereopsis (see Stimuli: Static RDSs section).

Figure 6

Anisotropic upper depth limits in owls O1 and O2. (A) $d_{\text{max}}$ determined from static displays of polarized RDSs. Each of the four panels shows a graphical representation of 6 successive staircase runs (black lines) produced by forced choice "hole" vs. "step" judgments. $d_{\text{max}}$ values for negative or positive disparity were derived separately from descending (shaded areas) or ascending staircases (non-shaded areas) starting at $-55$ or $+5$ arcmin, respectively. For example, a "back" response made the target disparity in the following trial more positive ($+5$ arcmin), whereas each "front" response made the target disparity for the following trial less positive ($-5$ arcmin). A single staircase ended (open rectangles) after 7 reversals, the direction wherein a change in disparity size is reversed. $d_{\text{max}}$ was calculated by averaging over a series of six staircases (dashed lines), whereby the first 2 reversals of each staircase were discarded and the remaining 5 reversals were averaged. To ensure stimulus control, catch-trial stimuli were presented 50% of the time. Catch-trial performance was high, above 85% correct; data not shown. Notice the severe anisotropy in the owls' upper disparity limits to positive ($+45$ arcmin) and negative ($-90$ arcmin) disparity-based RDSs. (B) As in (A) but now for $d_{\text{max}}$ determined from dynamic displays of polarized RDSs.

Procedure

$d_{\text{max}}$ (method of limits)

Upper depth limits for negative and positive center disparities were determined separately using a simple 1-up/1-down staircase procedure (converging to a probability correct of 0.50). Figure 6 provides a detailed description of
how this staircase procedure was implemented. To assess stimulus control, suprathreshold α RDS catch-trial stimuli were presented 50% of the time.

Results/discussion

In Figure 6, the individual graphs show that the upper depth limits ($d_{\text{max}} \pm [95\% \text{ confidence interval}]$ arcmin) determined from negative disparities in static RDSs (O1: 99.5 [97.6, 101.3] arcmin; O2: 98.8 [95.6, 101.4] arcmin) and dynamic RDSs (O1: 94.0 [92.0, 96.1] arcmin; O2: 97.1 [94.9, 99.5] arcmin) were, on average, 2 times larger than determined from positive disparities in static RDSs (O1: 45.7 [43.6, 47.9] arcmin; O2: 48.7 [46.3, 50.9] arcmin) and dynamic RDSs (O1: 43.8 [41.3, 46.3] arcmin; O2: 47.7 [44.3, 50.6] arcmin). Notably, this anisotropic $d_{\text{max}}$ was present irrespective of whether the RDSs were displayed statically or dynamically.

The extent of the difference in the owl’s ability to identify depth relationships from large negative and positive disparities along with the actual $d_{\text{max}}$ values that I obtained (≈100 arcmin for negative (crossed) disparities and ≈46 arcmin for positive (uncrossed) disparities) is in close quantitative agreement with the averaged values as obtained by Erkelens (1988) for human subjects under stabilized conditions of vergence.

Thus, although my measurements do not allow to distinguish between the disparity magnitudes at which differences in depth can still be perceived accurately and diplopia (double vision) begins or, alternatively, fusion is lost, it is likely that the upper depth limits determined here reflect the owl’s actual fusional limits. Note also that diplopia does not necessarily exist for RDSs. No matter how large the disparity, an RDS will not appear diplopic in the sense of being perceived with twice the dot density (Tyler, 1991).

Irrespective of whether the values obtained here reflect the owl’s true fusional limit or not, $d_{\text{max}}$ extended up to 100 arcmin. Thus, given that $d_{\text{max}}$ strongly depends on stimulus size, viewing duration, and random-dot density (which were in this case 2.5°, <1 s, and 50%, respectively), there is quantitative agreement with patent stereopsis in humans (Glennerster, 1998; Patterson & Martin, 1992; Tyler, 1991). This kind of coarse stereopsis refers to depth perception from binocular stimuli ranging from 1.0° up to 6.6° in size, with either fused or double, retinal images.

**Figure continuity**

Here, I establish if stereopsis can help owls determine what pieces of a visual scene actually compose the individual objects to be identified. To make this point, I conceived an experiment to test the gestalt principle of “figural continuity” or “amodal completion” (Kanizsa, 1979). It is derived after Nakayama’s figure–ground tests on surface representation in human vision (Nakayama, 1996; Nakayama, Shimojo, & Silverman, 1989).

If amodal completion takes place, then discontinuous surfaces will be seen as single object even when these discontinuities are caused by the presence of occluding surfaces. Consider, for instance, the black areas in the upper stereogram of Figure 7A; these are typically seen as a single object: the number 0. In contrast, the black areas in the lower stereogram of Figure 7A are seen as two disconnected U-shaped objects. Notice that in the upper stereogram the disparate random-dot region in the center is seen closer to the observer, whereas in the lower stereogram the same random-dot region is seen further away from the observer, when viewed under cross-fusion. Since the two stereograms only differ in the sign of disparity, the 2D perceptual organization of the two black U-shaped surfaces necessarily depends on the prior 3D interpretation of the entire scene.

In contrast to the problem of amodal completion—surface completion behind an occluder—we are sometimes faced with the problem to infer the existence of contours and surfaces in front of other surfaces, as might occur due to camouflage or poor illumination. Kanizsa (1979) described this perceptual puzzle as examples of modal or visible completions (for a review, see Nakayama, 1996; Nakayama et al., 1989). An example of modal completion is shown in Figure 8A (adapted after Nakayama, 1996; see p. 17). Observe that there is little tendency to see the cat in front of the rabbit or vice versa, when we view the half-images of top and bottom row stereograms monocularly. When fused as stereograms, however, their front–back depth relationship is immediately apparent and is reversed when the sign of relative disparity between the cat and rabbit is changed. Thus, disparity appears to dictate the depth placement of the rabbit relative to the cat. Most notably, this autonomous process is accompanied by the creation of subjective contours at the boundaries where either the cat silhouette overlaps that of the rabbit (upper stereogram) or vice versa (lower stereogram). Yet, how do owls interpret the ambiguous silhouettes of Figure 7A?

**Stimuli used to test for amodal completion**

A black number 0 ($9^\circ \times 9^\circ$) was embedded in a ($16^\circ \times 16^\circ$) static RDS (dot size: $1 \times 1$ arcmin; 50% dot density). The middle portion ($9^\circ \times 3^\circ$) of the 0 shape was excised. It was replaced by a set of random dots (dot size: $1 \times 1$ arcmin; 50% dot density) having either negative or positive disparity relative to the non-disparate surrounding random dots. Pictorial examples of these stereograms (along with their half-images) are shown in Figure 7A.

---

[Downloaded From: http://arvojournals.org/ on 10/29/2015]
Prior to exposure to the above-described stereoscopic stimuli, the owls were trained with flat, non-disparate stimuli (Figure 7B). Notably, these stimuli are identical to the ones depicted in Figure 7A except that the middle portion of the random dots was either removed (see the four exemplars on the left side, Figure 7B) or made non-disparate (see the four exemplars on the right side, Figure 7B).

In addition, I constructed control stereograms with irregular-shaped occluders (Figure 7D). The occluded region occupied ≈2 times the surface area compared to the one depicted in Figure 7A (upper row). Its shape and orientation was changed in a random fashion after every trial.

Stimuli used to test for modal completion

To create stereograms known to induce modal completion in humans, I drew a cat and a rabbit-like 2D silhouette (top row, Figure 8A). These flat, black painted silhouettes...
were approximately 12° × 18° in size. The surrounding, white background (0.36 cd m⁻²) was 25° × 25° in size.

The half-images of the stereograms needed to test for modal completion contained both 2D silhouettes laid on top of each other, with a small horizontal misalignment. Depending on the sign (i.e., direction) and magnitude of the horizontal misalignment, the cat would be seen in front of the rabbit (top row, Figure 8A) or vice versa (bottom row, Figure 8A). The disparity magnitude equaled +16 or −16 arcmin.

Procedure

**Amodal completion-based transitive inference testing**

The owls were subjected to a two-phase behavioral procedure. First, training took place with a 2D shape classification task under conditional reinforcement. Here, the owls had to indicate the presence of a continuous, 0-shaped object (left side, Figure 7B) or, alternatively, the presence of two disconnected, U-shaped objects (right side, Figure 7B), irrespective of their orientation. Second, during testing, the owls were rewarded after each trial. However, now, the random dots, separating the 0-shaped object into two U-shaped objects, could have a disparity other than zero; examples of which are shown in Figure 7A.

Threshold performance was determined using a forced choice back–front task in combination with a randomized block design, identical to the one described in the Stimuli (Static RDSs) section of Experiment 3: Acute stereopsis. The RDS’s occluder disparity was chosen from 11 possible values corresponding to −7.4, −4, −3, −2, −1, 0, 1, 2, 3, 4, or 7.4 arcmin. To assess stimulus control, suprathreshold α RDSs catch-trial stimuli were presented 20% of the time (N = 110 trials). For the parameterization of the fitted performance function, see the Psychometric functions section.

**Modal completion-based transitive inference testing**

The owls were subjected to a two-phase behavioral procedure. First, training took place with a 2D shape classification task under conditional reinforcement. Here, the owls had to indicate the presence of a cat or, alternatively, a rabbit-like silhouette (top row, Figure 8A). Second, during testing, the owls were rewarded after each trial. However, now, the cat and rabbit silhouettes were superimposed and presented as a stereogram where either the cat was seen in front of the rabbit (top row, Figure 8A) or, alternatively, the rabbit in front of the cat (bottom row, Figure 8A).

**Results/discussion**

**Amodal completion**

The owl’s ability to recognize the presence of a continuous 0-shaped object was assessed from the psychometric functions as shown in Figure 7C. Here, the probability of amodal completion—seeing a continuous 0-shaped object—is plotted as a function of the disparity of the “occluding” random dots. The shape of the psychometric curves confirms that the sign and magnitude of binocular disparity determines if a 2D surface remains segmented and separated from its neighbor (disparity ≥ 0) or, alternatively, if it becomes grouped (disparity < 0). First, in both owls, amodal completion only occurred when the discontinuous black surfaces were separated by a negative disparity-defined random-dot surface. Second, in both owls, the smallest possible relative disparity (σ ± 95% confidence interval arcmin) evoking amodal completion is in close agreement with the owls’ stereoaucity as determined from static RDS-based psychometric curves of Figure 4A (O1: 2.6 [2.0, 3.1] arcmin; O2: 2.5 [2.1, 2.9] arcmin). Finally, none of the owls displayed a response bias (μ ± [95% confidence interval] arcsec) deviating significantly from zero (O1: −36 [−42, +11] arcsec; O2: +7 [−14, +31] arcsec).

These data are all the more remarkable, because the owls were initially trained to select the “front” response bar when exposed to relative disparities smaller than zero (“step”, Figures 1A–1B). In marked contrast, here the owls choose the “back” response bar only in case of suprathreshold, negative disparity-defined stimuli. Importantly, in both owls, performance on the catch trials—containing suprathreshold α RDSs (as defined in Figure 3A)—was above 91% correct.

As a final control, I constructed RDSs with irregular-shaped occluder; examples of which are shown in Figure 7D. If only disparity information is responsible for amodal completion, then discriminative performance should not be dependent on the (surface) size, shape, position, and orientation of the occluder. High performance levels were obtained for both owls (O1: N = 100 trials, \( P_{\text{bin}} (X = 81\%) < 0.001;\) O2: \( N = 100\) trials, \( P_{\text{bin}} (X = 83\%) < 0.0001\) when exposed to these novel, randomly interleaved (suprathreshold: +3 or −3 arcmin) RDSs, under unconditional reinforcement. Most notably, negative disparity-defined stimuli were perceived to differ systematically from the positive disparity-defined stimuli [Fisher’s Exact Test, two-tailed, O1: \( \chi^2(1) = 40.9, p < 0.001;\) O2: \( \chi^2(1) = 46.6, p < 0.001\)]. Thus, it appears as if the owls apply a strategy in which disparity-defined amodal completion plays a dominant role in figure–ground segmentation. Taken together, the sign and amplitude of the relative disparity between 2D surfaces can play a decisive role in both border ownership and surface linkage of the owls’ visual perception.

**Modal completion**

To provide evidence for modal completion, owls O1 and O2 were trained to discriminate between the silhouettes of a rabbit and cat (learned categories, Figure 8A).
Hereafter, both owls were exposed—under unconditional reinforcement and without additional training—to the stereograms with superimposed cat–rabbit silhouettes (tested categories, Figure 8A). In 98% of all trials (N = 200), the owls associated the upper stereogram with the silhouette of a cat or, alternatively, the bottom stereogram with the silhouette of a rabbit (Figure 8B). Evidently, owls are able to bind contours—modal completion—that are incomplete due to an “inferred” occluding surface.

Critical to the design of the stereograms used to test for modal completion was the occurrence of partial occlusion and disparity cues bounded by illusory contours that are induced by familiar shapes, in this case the learned cat and rabbit silhouettes. To my knowledge, this is the first behavioral evidence outside the human (e.g., see Burge, Fowlkes, & Banks, 2010) and non-human primate domains (e.g., see Bakin, Nakayama, & Gilbert, 2000; Bushnell, Harding, Kosai, & Pasupathy, 2011), showing that information about depth in one part of the visual scene propagates to other parts of the scene while at the same time influencing 2D form perception (for a review, see Anderson, 1998; Nakayama et al., 1995). This is not unexpected since occlusion in particular poses a ubiquitous problem, given the multiplicity of objects in the world and the loss of one spatial dimension (3D to 2D) during image projection (Singh, 2004).

**Figure 8.** Disparity-based modal completion of ambiguous surfaces in owls O1 and O2. (A) A cat-like and a rabbit-like silhouette (learned categories) along with two stereograms (tested categories) where the two silhouettes are superimposed. As such, the depth relationship between the cat and rabbit silhouettes of the respective half-images is ambiguous. Stereoscopic viewing of the half-images, however, can result in a dramatic perceptual change. In humans, cross-fusion of the top stereogram reveals the cat in front of the rabbit, whereas in the opposite stereo case (bottom stereogram) the rabbit is seen in front of the cat. Apparently, disparity information can be used to bound contours—modal completion—that are incomplete because of an “inferred” occluding surface. (B) The bars show the mean percentage of “cat” choices (left bars: O1: 100%; O2: 99%) when the owls were presented with “cat in front” stereogram or “rabbit” choices (right bars: O1: 99%; O2: 100%) when presented with the “rabbit in front” stereogram. Each bar represents 10 transfer sessions. N denotes number of trials. The gray dashed lines at 50% correct mark the performance level expected when the owl would be guessing.

The results of psychophysical experiments with RDSs can be contaminated by (1) cues originating from the equipment, (2) the technique used to create RDSs, or (3) perceptually induced cues such as false motion or monocular occlusion. These unwanted cues might not be obvious to human observers (Cobo-Lewis, 1996; Collewijn, Steinman, Erkelens, & Regan, 1991; Hoffman, Girshick, Akeley, & Banks, 2008; Tyler, 1991). Note, however, that although false motion and monocular occlusion are not necessarily artifacts of the stimulus-presenting method, it is not known whether or not these cues can be used by the owl visual system to perceive depth.

### Artifacts and stimulus-presenting method

To ensure contamination-free stimuli and stereopsis-driven performance, I worked out two controls: (1) mono viewing (Fox et al., 1977) and (2) an alternative method for presenting dichoptic stimuli: anaglyph RDSs (Julesz, 1971), respectively. Mono viewing was tested by placing filters of the same polarization, or color, in front of their eyes. The application of this procedure permits only one half-image to stimulate both eyes. When a single technique is applied to create RDSs, however, it is impossible to know whether the resulting psychophysical data are due to perception itself or to response processes inherent in the RDS technique (Wagner, 1985). Thus, when discriminative...
behavior is purely based on stereopsis, then performance should only be affected by mono viewing and not by a change in the stimulus-generating method.

Mono viewing sessions were of two types. Either the left or, alternatively, the right half-image of an RDS was made visible to both eyes. This control was performed with both polarized and anaglyph RDSs. A single experiment consisted of a sequence of five sessions (of 12 trials each: 6 “steps” plus 6 “holes”) alternating between the stereoscopic and mono viewing conditions. The depth stimuli were identical to the a RDSs of the depth-ordering stereograms (defined in Figure 3A).

Discriminative performance was significantly changed by mono viewing but not by the alternative RDS-generating method (Figure 9A, trained polarized vs. non-trained anaglyph RDSs). Most importantly, the latter occurred despite the fact that the owls were trained with polarized RDSs alone. A Fisher exact test, \( \chi^2 \), as applied on the stimulus–response relationships of the pooled data (including both polarized and anaglyph RDSs), showed that hole- and step-shaped RDSs were only perceived to differ when viewed binocularly (O1: \( N = 480 \) trials, \( \chi^2(1) = 32, p < 0.001 \); O2: \( N = 480 \) trials, \( \chi^2(1) = 23, p < 0.001 \)). Conversely, stereo performance fell to chance levels when the filters were of the same polarization or color, indicating that the amount of luminance captured by the blocked eye was negligible. Evidently, performance was not aided by artifacts in the RDSs or the stimulus-presenting technique (polarization vs. anaglyph) and was solely determined by the owls’ ability to detect differences in absolute, horizontal binocular disparity.

### False motion

My interpretation of the data has assumed that the owls viewed the RDSs motionless. However, barn owls make spontaneous head movements that mainly occur in the horizontal plane while exploring a visual scene (van der Willigen et al., 2002). This so-called peering behavior can raise a potential complicating factor, referred to as false motion or pseudo-parallax. In particular, human observers that peer while viewing RDSs experience a distortion suggestive of motion parallax (a monocular depth cue), even though there are no parallax cues (Tyler, 1974). That is, the lack of relative motion within the stimulus itself. Specifically, lateral distortions of the stereoscopic central region relative to its surrounding background is most likely due to the visual system’s temporary failure to reject the implications of the absence of any motion parallax when the head is moved sideways (Howard & Rogers, 2002). Thus, false motion can be exploited to aid depth perception from RDSs (see also Hoffman et al., 2008).

I do not know if owls experience false motion. Even if they do, owls might not (mis)interpret the cues involved in the same way as humans might. Still, it is quite possible that owls take advantage of peering movements to produce pseudo-parallax that help in their judgment of RDSs.

To establish if reliable performance was accompanied by peering movements, head position was measured when the owls were presented with novel, asymmetric versions of the RDSs used in the depth-ordering task. Specifically, the stereoscopic central regions of the \( \beta \) and \( \gamma \) depth configurations (as defined in Figure 3A) were reduced to \( 50 \times 50 \) dots and could occupy one of four possible locations, each specified by the center coordinates of the four quadrants of the surrounding ground. In addition, the amount of disparity was reduced by a factor of 0.625. My objective was to evoke systematic lateral (X-axis, Figure 1Aiii) changes in head position just after stimulus onset by (1) making the “hole”/“step” judgments more difficult and (2) by presenting the disparate random dot away from the line of view at stimulus onset.

Both owls exhibited high performance levels significantly deviating from random responding (O1: \( N = 40 \) trials, \( P_{\text{bin}}(X = 90\%) < 0.0001 \); O2: \( N = 40 \) trials, \( P_{\text{bin}}(X = 95\%) < 0.0001 \)). Systematic changes in head position just after stimulus onset, <700 ms, were not observed, as can be seen from the individual (horizontal) head movement traces of Figure 9B. Visual inspection of the \( x \), \( y \), and \( z \) positions of each individual head movement trace, 80 in total, reveals that deliberate head movements were produced only to operate the response bars, except for one trial in which a horizontal peering movement occurred prior to the conditioned response (Figure 9B, arrow). Evidently, owls, like humans, do not require false motion to perceive relative depth from RDSs.

### Monocular occlusion

The technique used throughout this study for demonstrating disparity-based figure–ground segmentation exploits an assumed symmetry about depth order. By simply interchanging the left and right eyes’ images of a stereogram, the stereoscopic form is not altered, but the sign of disparity is inverted, causing a depth inversion; see, e.g., the stereograms of Figure 7A. As we have seen, this rather simple experimental manipulation provides an effective means of assessing the contribution of disparity to stereopsis. However, when the half-images of a stereogram are interchanged, not only the disparity sign is inverted, it also changes the occlusion relationships. To see this, I study how the RDSs of Figure 1Aii is constructed. Notice that the unpairable gaps as left behind by the shift of the center square are only visible in one of the half-images. It is not hard to imagine that when the half-images are interchanged, the relationship between these so-called monocular occlusions is also interchanged (see also Anderson, 1998).
The aforementioned technical considerations about the peculiarities of interchanging half-images leave open the possibility that occlusion alone may be sufficient to evoke figure–ground segmentation. This is not without precedence since human stereovision is susceptible to systematic changes in figure–ground segmentation that cannot be understood on the basis of disparity processing alone. In particular, Kaye (1978) was the first to describe the utility of monocular regions for depth perception (for review, see Harris & Wilcox, 2009).

Moreover, while disparity provides relative depth information about surfaces visible to both eyes, half-occlusions (Belhumeur & Mumford, 1992) provide monocular information—seen by one eye alone—that can be used geometrically to segment visual space into figure and ground. Half-occlusions, thus, refer to unpairable image features that arise at monocular (or half-occluded) contour regions.

For the case of depth created by monocular occlusion in human vision, consider the stereograms of Figure 9C. Here, neither the left nor the right contour of the “disparate” black central region has a corresponding luminance-contrast contour in the opposite half-image. Thus, the figure–ground depth impressions created by cross-fusion of the half-images—top stereogram: rectangle in front of a slit; bottom stereogram: rectangle behind.
a slit—are due to monocular occlusion alone. Yet, how do owls interpret such half-occluded stereograms?

Owls O1 and O2, naive to the half-occlusion stereograms of Figure 9C, immediately generalized to the perceptual problem of monocular occlusion-induced depth perception (O1: N = 100 trials, Pbin(X = 95%) < 0.0001; O2: N = 100 trials, Pbin(X = 92%) < 0.0001), under unconditional reinforcement (Figure 9D). In other words, owl stereo vision propagates depth information inferred from half-occluded regions to other parts of the visual scene that do not “own” enough monocular visible borders by themselves to form separate surfaces.

Thus, owl stereovision, like human stereovision, appears to have explicit knowledge about the geometric origins of half-occlusions, because this information can be used reliably—without prior learning—to determine matching figure–ground depth/contour relationships.

General discussion

Despite considerable attention within the neurosciences, there is an unresolved contradiction between the widely shared view that human stereopsis is exceedingly effective at recovering depth relationships through the sole detection of disparities (Georgieva et al., 2009; Parker, 2007) and its possible biological significance in vertebrate vision (Anderson, 1998; Burge et al., 2010; Collett, 1996; Farel, 2006; Glennerster et al., 2006; Heesy, 2009; Julesz, 1994; McKee, Levi, & Bowne, 1990; McKee, Watamaniuk, Harris, Smallman, & Taylor, 1997; Pettigrew, 1986; Ponce & Born, 2008; Tyler, 2004a). At the heart of this contradiction lies Julesz’s (1964) RDS-based demonstration that disparity alone—the strongest metric depth cue available (Todd & Norman, 2003)—can evoke vivid impressions of depth (Figure 1Aii), while at the same time humans are generally incapable of making reliable judgments about the metric of 3D shape (for review, see Todd, 2004).

Thus, the question arises: “What perceptual problem does stereopsis solve given that disparity sensitivity does not necessarily result into veridical impressions of 3D space?” Here, I make the case that this question can be answered, in part, by drawing together RDS-based stereo performances from both owls and humans.

Overall, I identified six perceptual biases of RDS-based stereopsis that owls share with humans: (1) Owls recognize 3D shape from RDSs in the absence of monocular depth and familiarity cues and of all binocular depth cues except for disparity. Thus, stereopsis provides a computational strategy with which observers can avoid feature extraction until information from both eyes has converged. (2) Owls rely on relative, rather than absolute, disparity. Like humans, owls are poor at detecting global changes in absolute disparity. This eliminates the fovea as a key reference point, making disparity sensitivity independent from convergence eye movements. Thus, stereopsis in itself does not provide absolute depth information. (3) Owls enjoy the same improvement in angular resolution due to binocular processing as humans do when taking into account both the MAR and IPS of the respective visual systems. Thus, like humans, owls exhibit hyperacuity, where stereo thresholds are significantly lower than predicted from the anatomical resolving power of the eye. This makes stereopsis well suited to detect the mere presence of minute changes in depth. (4, 5) Owls have upper depth limits within the domain of human patent
stereopsis and can perceive disparity-defined form in less than 50 ms when (i) convergence movements of the eyes, (ii) proprioceptive cues, and (iii) top-down influences are by and large eliminated. The relative unimportance of the convergence angle and large disparities may be inconsequential for the owl, with its short talons, its ability to fly, and little or no convergence capability (for review, see Martin, 1990), but this is certainly not true for humans with their fast and highly sensitive convergence capabilities (Masson et al., 1997) and long, agile forelimbs that can easily reach for nearby objects (Heesy, 2009). Because of these limitations, stereopsis seems not geared to control goal-directed movements, but instead, it allows for the fast detection of minute discontinuities in depth. (6) Owls, like humans, can exploit disparity information to link spatially separated surfaces that belong to a single, partially occluded object (amodal completion). Thus, binocular comparisons aid in distinguishing the boundaries of tangible objects from those created by occlusion.

The cross-species evidence highlighted in this study suggests that the full range of perceptual biases associated with stereopsis cannot be readily explained in terms of depth perception alone. For example, the lack of metrical precision in disparity-based depth perception makes it difficult to conceive how stereopsis in early primates could have been important for the transformation of arms and limbs into delicate grasping organs. Without denying disparity’s role in perceiving depth, however, my data can be easily rationalized in terms of 2D perceptual organization. That is, disparity essentially tells owls, like humans, about relative depth rather than absolute distance and is useful, therefore, not so much in controlling goal-directed movements, as in recognition (Carpenter, 1991; Collett, 1996; Nakayama, 2005).

Acknowledgments

The author gratefully acknowledges the support of Dr. Hermann Wagner who made this work possible. All animal work was performed at the Department of Zoology and Animal Physiology, Rheinisch-Westfälische Technische Hochschule (RWTH), Aachen, Germany.

The German Research Foundation (http://www.dfg.de) supported the animal research. The writing of this paper was made possible by the Neurocognition Program of the Netherlands (051.04.022), as well as a VICI Grant to Dr. A.J. van Opstal (ALW 865.05.003) within the Earth and Life Sciences of NWO (http://www.nwo.nl/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Commercial relationships: none.

Corresponding author: Robert F. van der Willigen.

Email: r.vanderwilligen@donders.ru.nl.

Address: Department of Biophysics, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Geert Grooteplein N 21, 6525 EZ Nijmegen, The Netherlands.

References


Bravo, H., & Pettigrew, J. D. (1981). The distribution of neurons projecting from the retina and visual cortex to the thalamus and tectum opticum of the barn owl, Tyto alba, and the burrowing owl, Speotyto
cunicularia. Journal of Comparative Neurology, 199, 419–441.


Collins, E. T. (1921). Changes in the visual organs correlated with the adoption of arboreal life and with the assumption of the erect posture. Transactions of the Ophthalmological Society, 41, 10–90.


