Voluntarily controlled bi-stable slant perception of real and photographed surfaces

Raymond van Ee[1*], Gunta Krumina[2], Sylvia Pont[1] and Sanne van der Ven[1]

[1]*Utrecht University, Helmholtz Institute, 3584 CC Utrecht, The Netherlands*
[2]*Department of Optometry and Vision Science, University of Latvia, Riga, 1063 LV, Latvia*

We have quantified voluntarily selected perceived slant of real trapezoidal surfaces (a 'reverse-perspective' scene) and their photographed counterparts (pictorial space). The surfaces were slanted about the vertical axis and observers estimated slant relative to the frontal plane. We were particularly interested in those cases in which binocular disparity and monocular perspective provided conflicting slant information. We varied the monocularly and binocularly specified surface slants independently across stimulus presentations. To eliminate texture and shading cues we used sand-blasted aluminium trapezoidal surfaces illuminated from all directions. When disparity-specified slant and perspective-specified slant were conflicting, observers were able to perceive the surfaces in two ways: they perceived either a trapezoid or a rectangle. Our main finding is twofold. First, when subjects chose to perceive the trapezoid, the slant estimates followed the disparity-predicted slant with only a slight underestimation, as if they selected a pure binocular representation of slant governed only by disparity. Second, when subjects chose to perceive the rectangle their estimates for real surfaces were similar to those for photographed surfaces, as if they selected a representation of slant governed by perspective foreshortening.

Keywords: reverse perspective; perceptual bi-stability; voluntary control; perceived surface orientation; pictorial space

1. INTRODUCTION

In the phenomenon of visual bi-stability a constant retinal image produces a changing percept. It is an interesting phenomenon because it raises the possibility of having two states in processing that are modulated by the observer's assumptions about the world rather than by the stimulus. Perceptual bi-stability has been successfully used to study visual processing, including some aspects of visual awareness (review in Blake & Logothetis 2002). Numerous reports show that the perceptual alternation frequency in bi-stability is, although to a limited extent, under voluntary control, making perceptual bi-stability an even more scientifically appealing phenomenon (review in van Ee *et al.* 2005). Further, perceptual bi-stability is interesting because it challenges theories that relate the quantitative aspects of perceived depth to the available depth cues.

To study how voluntarily selected percepts are related to the quantitative aspects of stimuli, we recently developed a 'slant rivalry' paradigm which capitalizes on the distinct binocular and monocular depth information in an image (van Ee *et al.* 2002; van Ee 2005). Binocular disparities arise because our eyes view a scene from slightly different positions. These disparities enable us to perceive the 3D layout. Monocular cues can also be sufficient to recover the 3D layout. For example, linear perspective is a powerful cue for surface orientation. The integration of perspective and disparity has been the subject of several studies (reviewed in Howard & Rogers 2002). However, the

bi-stability that can be created when the monocular and binocular cues in a scene specify opposite depth information has attracted little interest, and few studies have modelled the quantitative aspects of this phenomenon. We recently developed a Bayesian model for the quantitative aspects of bi-stability in perceived slant for many combinations of disparity- and perspective-specified slants. Although there are fundamental differences between observers we are able to explain the metrical aspects of perceived slant on the basis of the relative likelihood of both perspective and disparity slant information, combined with prior assumptions about the shape and orientation of objects (van Ee *et al.* 2003).

Is the metrical relationship between perceived slant and both the perspective and the disparity signals, which we found previously, a curiosity of stimuli produced by stereograms on a monitor? This is an important issue because we know from the literature that real 3D stimuli can play a distinct role in perceived depth (e.g. Frisby *et al.* 1995; van Ee *et al.* 1999) presumably because conflicting signals that inform the subject about the flatness of a monitor are not present in real 3D stimuli. van Ee *et al.* (1999) showed that the classical stereoscopic slant contrast effect (Werner 1937) is peculiar to stereograms. They developed a slant-perception theory in which the most reliable slant signal would get the most weight. Their theory predicted a so far undescribed and curious effect: namely that the conventional direction of slant contrast would reverse if the surface slant is governed by perspective signals rather than by disparity signals. Using both stereograms on a monitor and real wooden plane stimuli they confirmed their theory's predictions, which led them to conclude that slant contrast is nothing more than a by-product of the visual system's

reconciliation of conflicting information while it attempts to determine surface slant. Frisby *et al.* (1995) used real and stereogram stimuli to study the role of blur cues as a factor in studies involving conflicts between disparity and perspective cues. They concluded '*Beware drawing firm conclusions from stereograms about the pattern of cue integration that can be expected when real objects are being viewed*'.

To explore metrical aspects of the mental process that underlies perceptual alternations we examined perceived slant when slant rivalry was produced by real objects. An example of a real plane stimulus that consists of conflicting perspective and disparity cues is the well-known Ames trapezoid (Ames 1951). Another striking example of depth inversion is the hollow relief mask, which can be seen in reversed perspective (Yellott & Kaiwi 1979). One of the most interesting (and enjoyable) examples of displays that can be used to study perspective-disparity cue integration is the reverse-perspective paintings on 3D canvas by Patrick Hughes (see Slyce 1998 for many paintings). Figure 1*a* illustrates how to construct a simple 3D replica of a reverse-perspective scene (Wade & Hughes 1999). Reverse-perspective scenes capitalize on the perceptual alternations that Ernst Mach observed *ca.* 150 years ago when he folded visiting cards and placed them so that they were illuminated more from one side than from the other side (Mach 1866). The reverse-perspective scenes are attractive, both in art and in research (Cook *et al.* 2002; Papathomas 2000, 2002), because they bring about a conflict between depth specified by perspective and depth specified by disparity. The foreshortening (linear perspective) of the portrayed door in figure 1*b* specifies that the door's right side is receding in depth. Because, in reality, the door's right side is protruding, the disparity-specified slant is opposite to the perspective-specified slant. The stereogram in figure 1*c* illustrates the phenomenon of perceptual alternations in slant rivalry. After fusion of the stereogram two relatively stable percepts can be distinguished. In the first percept the right side of the door appears further away (it is perceived as a normal slanted rectangular door). In the other percept, the left side of the door is further away (it is perceived in reverse perspective: as a trapezoidal door with the near-edge shorter than the far-edge). Each percept can be selected and maintained at will in a relatively controlled way. Which of the two percepts dominates depends on the viewing distance (Papathomas 2000, 2002). Binocular disparity is dominant for short viewing distances and monocular cues are more important at larger viewing distances.

We initially set out to study depth cue integration by using painted reverse-perspective scenes (say, one of Hughes's paintings). However, realistic scenes contain various cues to depth such as shading, texture and occlusion that complicate a systematic scientific experiment. We therefore used sand-blasted aluminium trapezoidal stimuli in which disparity and perspective specified different slants. Figure 2 shows examples of the reverse-perspective slant stimuli used in this study. Slant refers to surface rotation about a vertical axis through the centre of the stimulus (Gillam 1968). To eliminate texture cues the stimuli were sand-blasted homogeneously with very fine grain that could not be resolved at the viewing distance. To eliminate shading cues the stimuli were placed in a large booth and illuminated from all directions.
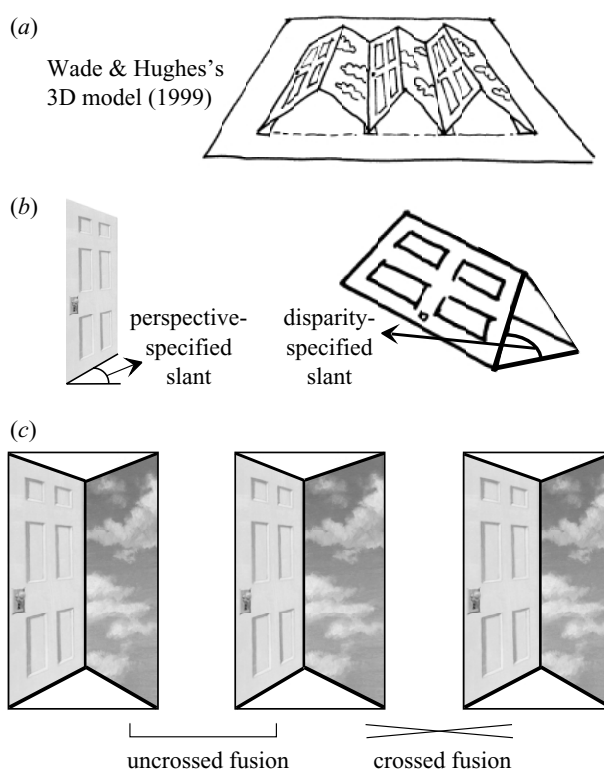
Figure 1. The reverse-perspective phenomenon. (*a*) How to construct a 3D replica of a reverse-perspective scene. The main characteristic that makes reverse-perspective scenes attractive, both in art and in research, is that those scenes bring about a conflict between the depth specified by perspective and the depth specified by disparity. (*b*) An explanation of the foreshortening (linear perspective) of the portrayed door that specifies that the door's right side is receding in depth. Because, in fact, the door's right side is protruding the disparity-specified slant is opposite to the perspective-specified slant. The stereogram in (*c*) illustrates what observers perceive when viewing one panel of the 3D replica. After fusion of the stereogram two relatively stable percepts can be distinguished. In the first percept, the door recedes in depth with its right side further away (it is perceived as a normal slanted rectangular door). In the other percept, the left side of the door is further away (it is perceived in reverse perspective: as a trapezoidal door with the near-edge shorter than the far-edge). Each of the percepts can be selected and maintained at will in a relatively controlled way. When the left two images are being fused in a crossed way (or the right two images in an uncrossed way), perspective and disparity specify similar slants and the observer perceives a single stable slanted rectangular door with its right side further away. Adapted from the 'Cloudy Doors' 3D model of Wade and Hughes (http://www.perceptionweb.com/perc0999/wade.html) with the permission of the authors and Pion Limited, London.

To explore perceptual bi-stability in further detail, we also examined the metrical aspects of perceived slant of photographs of the stimuli. Pictorial space refers to the 3D spatial impression obtained when one looks at 2D photographs (review in Koenderink & van Doorn 2003; or see Ellis *et al.* 1991). Photographs are usually viewed with two eyes. Because it is *a priori* not clear to what extent there is a difference between monocularly and binocularly estimated slant our observers performed their slant estimates under both viewing conditions. Several studies have examined

Figure 2. The stimuli used. Each of the depicted stimuli is a sand-blasted aluminium trapezoid for a specific combination of foreshortening and disparity. Some of the stimuli are both slanted and a little rotated relative to the black table on which they are lying. This gives rise to the apparent deformations on this picture.

pictorial space engendered by natural objects (e.g. van Doorn *et al.* 2001) or scenes (Hecht *et al.* 1999). We are specifically interested in perceived slant. Although some studies have addressed this issue (Rosinski *et al.* 1980; Kubovy 1986) systematic studies for different perspective-specified slants have not been reported.

2. MATERIAL AND METHODS

(a) *Experiment 1*

(i) *Stimuli and apparatus*

Figure 2 portrays the trapezoidal planes, each having a unique combination of disparity-specified and perspective-specified slant. The disparity-specified and the perspective-specified slants could both be $-70°$, $-50°$, $-25°$, $0°$, $25°$, $50°$ or $70°$, yielding a total of 49 (7×7) stimuli. The subtended horizontal angular size of the stimuli was always $7.8°$. The vertical size was always $7.8°$ at the location of the slant axis. However, the foreshortening was different for each stimulus. Figure 3 explains how differently slanted aluminium planes (different disparity-specified slants) could have the same perspective-specified slant (the same visual angles from the point midway between the eyes). We constructed only 25 aluminium trapezoidal planes because the other 24 planes were obtained by rotating them $180°$ about the viewing axis. For example, the combination disparity- and perspective-specified slant ($25, 25$) specified the opposite slant, namely ($-25, -25$) after rotating the aluminium plane.

The slanted planes were illuminated from all directions to ensure that there were no shading cues. The diffuse (Ganzfeld) illumination was produced by a large illumination booth (see figure 4). Each of the six sides of the booth consisted of a white opal glass pane subtending $1\,\text{m} \times 1\,\text{m}$. The panes were backlit by fluorescent tubes, resulting in spherical diffuse illumination in the centre of the booth. One side of the booth could be opened to place the slant stimuli in the booth. The stimuli were supported in the centre of the booth by a hidden rod extending to the back of the booth. The mounting device did not noticeably disturb the diffuse illumination. The slanted stimuli were viewed through a circular aperture in one of the side panels of the booth (figure 4*b*). The distance between the subject's eyes and the stimulus was 72 cm. The subject's chin was held by a chin cup. There was no fixation point. In front of the subject, at a distance of 25 cm, a binocularly visible 'slant-matching' bar could be rotated around the vertical axis by an electric motor controlled by a joystick.
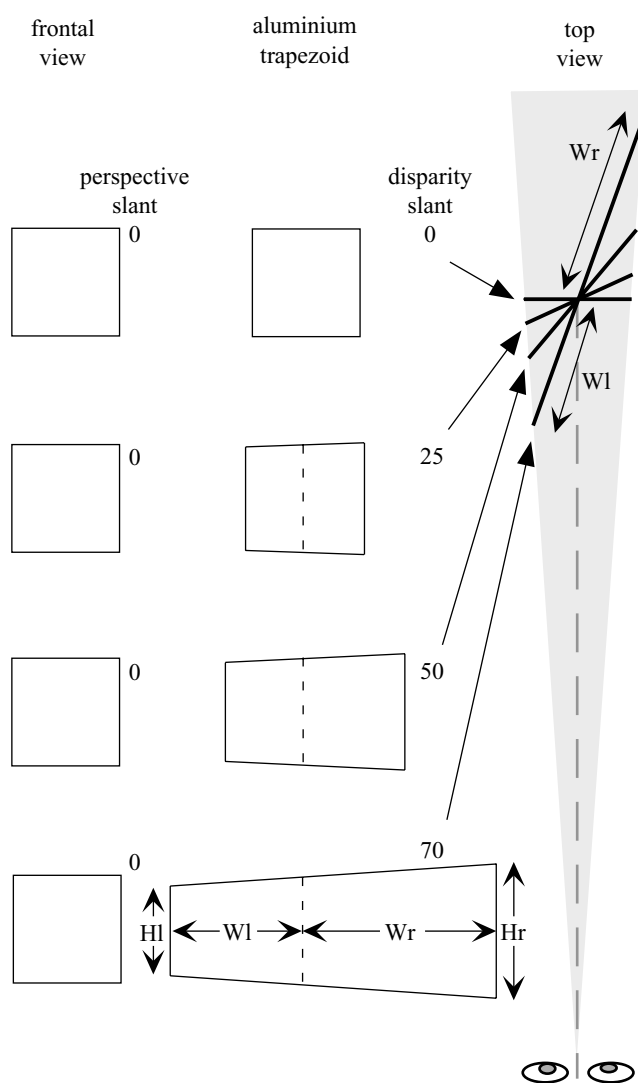

Figure 3. The geometry of reverse perspective. To create an aluminium trapezoid with disparity-specified slant and perspective-specified slant that differ from one another, we varied the stimulus heights and widths on its left (Hl, Wl) and right (Hr, Wr) sides. We show an example of a stimulus for which perspective specified zero slant (left column) under different disparity-specified slants (right column). We kept the horizontal visual angle that the trapezoid subtended (grey area) constant across stimuli. Note that this means that the Wl and the Wr of a stimulus are generally unequal.

Subjects could see this slant-matching bar and the stimulus at the same time without making head movements.

(ii) *Procedure and task*

The experimenter placed the slanted planes in the illumination booth while the viewing aperture was closed. The stimuli were presented in a random order. Subjects were told that ambiguous (bi-stability) and non-ambiguous (no bi-stability) stimuli would be presented and that the stimuli could be either trapezoidal or rectangular. Note that linear perspective information in an image can only be exploited by making assumptions about the orientations of the contours in the world that are being projected onto the image plane. In our study, perspective information was interpreted by assuming the object is rectangular (e.g. Clark *et al.* 1956; Reinhardt-Rutland 1990; van Ee *et al.* 2003). Subjects were instructed to voluntarily select either the trapezoid or the rectangle
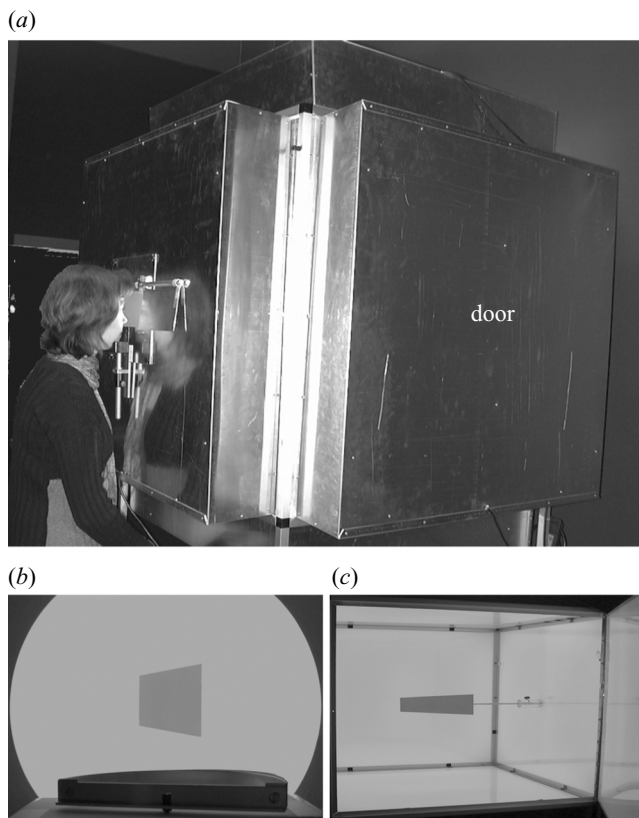
(a)

(b)   (c)

Figure 4. Experimental set-up. (a) The illumination booth. The subject views the stimulus through an aperture while he/she matches the slant of the rotatable device that is visible on the foreground (b). The booth has a door that can be opened to enable the experimenter to place the stimuli in the booth (c). Note that the depicted trapezoid recedes in depth with its right side further away (c). In fact, its slant is $70°$. However, the perspective-specified slant in (b) strongly indicates that the left side recedes in depth. In other words, we have here a reverse-perspective scene (compare with figure 1) under well-controlled visual conditions.

percept and to estimate the slant of the perceived surfaces. This was done first for the form that the subject initially perceived (say, a rectangle with its left side in front), and then, if the percept was bi-stable, for the second perceived slant (the trapezoid with its right side in front). When a subject was unable to experience bi-stability, only one slant setting was recorded. The viewing period was unlimited. Each trial block consisted of the above-described 49 stimuli. Each subject completed three trial blocks with binocular vision followed by two with monocular vision (always the left eye).

(iii) *Subjects*

Eight subjects with normal or corrected-to-normal vision participated. Their stereo vision was tested by a stereo-anomaly test of the ability to distinguish between crossed and uncrossed disparities (defined relative to the monitor) of magnitudes within a range of $-1°$ to $1°$, without the possibility that eye movements interfere (van Ee & Richards 2002). Subjects NK, MS and SV were excellent at distinguishing the signs and magnitudes of both the crossed and the uncrossed disparities. RR and TV were significantly above chance in perceiving crossed and uncrossed disparities without eye movements, and were excellent when eye movements were allowed. The sixth and seventh subjects, GK and SP, were able to correctly process the crossed disparities (within a range of $-1°$ to $0°$), but not the uncrossed disparities. These

subjects had to rely completely on eye movements to make correct disparity slant judgements. The eighth subject, MK, was invited to participate because her vision was dominated by monocular vision (caused by a pathological history). Prior to participation, the subjects were also tested for consistency in their responses when estimating the slants of both real and dichoptically presented planes.

(b) *Experiment 2*

(i) *Stimuli and apparatus*

The slanted aluminium stimuli that were presented in experiment 1 were photographed with a digital camera through the circular viewing aperture of the illumination booth. Figure 4b illustrates how a photographed stimulus looked to the subject (the matching bar visible in figure 4b was not visible). The RGB files were calibrated photometrically such that the minimum and the maximum RGB-levels represented black (measured with a black standard) and white (the background), respectively. We checked whether the geometrical properties of the stimulus were preserved in the pictures. If correct, we cut 1600×1600 images, which were saved with a resolution of 600 dpi ($6.77 \text{ cm} \times 6.77 \text{ cm}$). The calibrated, rotated and cut images were down sampled to 400×400 pixels, 72 dpi, and saved as RGB PICT files (no compression, 32 bits pixel^{-1}). These PICT-files were presented in random order on a LaCie (electron 22 blue IV) high-resolution monitor of $31.2° \times 22.8°$, such that the angular sizes of the stimuli were the same as they were when presented in the illumination booth. The viewing distance was 72 cm, as in experiment 1. The head was stabilized by a chin and forehead rest.

(ii) *Procedure and task*

We instructed subjects to estimate the slant of the plane displayed on the monitor. Subjects were asked to estimate only the slant of the perceived rectangle. In other words, we told subjects that the photographs were made of slanted rectangles, but not of slanted trapezoids. We asked for only one response (both for the binocular and the monocular viewing conditions) because in pilot experiments we found that it was uninformative to ask for an estimate of the slanted trapezoid. This is not to say that the subjects did not experience bi-stability. For each of the presented photographs subjects were always able to perceive the plane as unslanted (as specified by all cues except foreshortening): disparity always indicated zero slant.

The slant estimation procedure has been previously described in detail (van Ee *et al.* 2002). In short, after presentation of the stimulus, two lines were presented on the monitor. One of the lines was horizontal and the other line could be rotated about its centre. The horizontal line was fixed and represented a top view of the unslanted reference; the other line represented the top view of the perceived slanted surface. Subjects were instructed to match the angle between the rotatable line and the horizontal line to the two perceived slants. As in experiment 1, there were 25 different stimuli (photographs) of which 24 could be used in reverse orientation, amounting to 49 stimuli. Each subject completed three trial blocks with binocular vision followed by three with monocular vision. Out of the eight subjects who participated in experiment 1, five subjects were available for experiment 2: GK, MS, NK, SP and SV.

3. RESULTS

(a) *Experiment 1*

Subjects GK, MS, NK, SP and SV participated in experiments 1 and 2. For comparison, we present their data for the two experiments in the same graphs. Figure 5a depicts
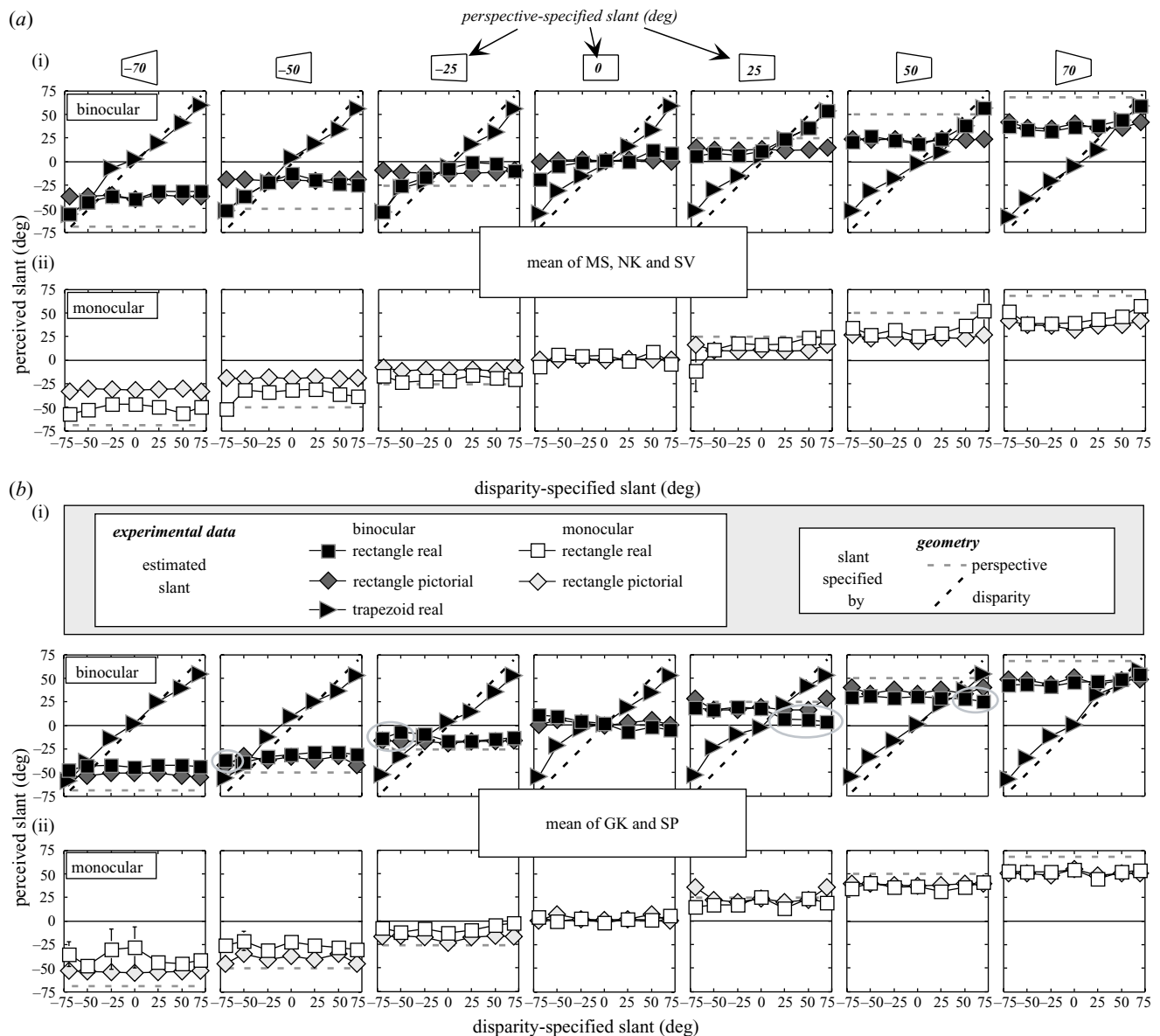
Figure 5. Data from experiments 1 and 2 for the subjects who completed both experiments. Perceived slant is plotted as a function of disparity-specified slant for a range of different perspective-specified slants. The trapezoid-shaped icons above the plots depict the perspective-specified slant. (*a*(i),*b*(i)) and ((*a*(ii),*b*(ii)) depict binocularly (filled symbols), and monocularly (open symbols) perceived slant, respectively. The slants that were geometrically present in the stimulus are represented by the dashed prediction lines. The data of experiment 1 are represented by the square and triangle symbols: subjects perceived either a slanted rectangular surface (squares) or a slanted trapezoid (triangles). The data of experiment 2 are represented by the diamond symbols: subjects perceived a slanted rectangular surface on the pictures. (*a*) The mean data of MS, NK and SV. (*b*) The mean data of GK and SP. Error bars, which are often smaller than the symbol, represent ±1 s.d. in the mean across the participating subjects.

the data of subjects MS, NK and SV, and figure 5*b* the data of subjects GK and SP. The data of these two groups of subjects are presented separately because, as we will see, they show interesting differences. Subjects TV, RR and MK participated in only experiment 1 (figure 6). All plots depict the mean perceived slants across subjects versus the disparity-specified slants. Each of the plots shows the data for a particular perspective-specified slant that is denoted by the trapezoid-shaped icon above the plots. The square symbols in all graphs indicate the estimated slant when the subjects perceived a slanted rectangular surface. The triangles indicate the estimated slant when they perceived a slanted trapezoid. Filled symbols denote binocularly estimated slant, open symbols denote the monocularly

estimated slant. The grey and black dashed lines denote the geometrically predicted slant based solely upon perspective and disparity, respectively.

We first address the results of subjects MS, NK and SV (figure 5*a*). Their plots for binocularly perceived slant (figure 5*a*(i)) can be roughly split into two domains: in the first domain subjects reported only one perceived slant. In this domain, slants derived from perspective and disparity were reconciled, producing a slant estimate somewhere between the two. Even if disparity and perspective specify identical slants, there is the often-reported slant underestimation (Howard & Rogers 2002). In the second domain, when disparity and perspective specified very dissonant slants, subjects experienced bi-stability and
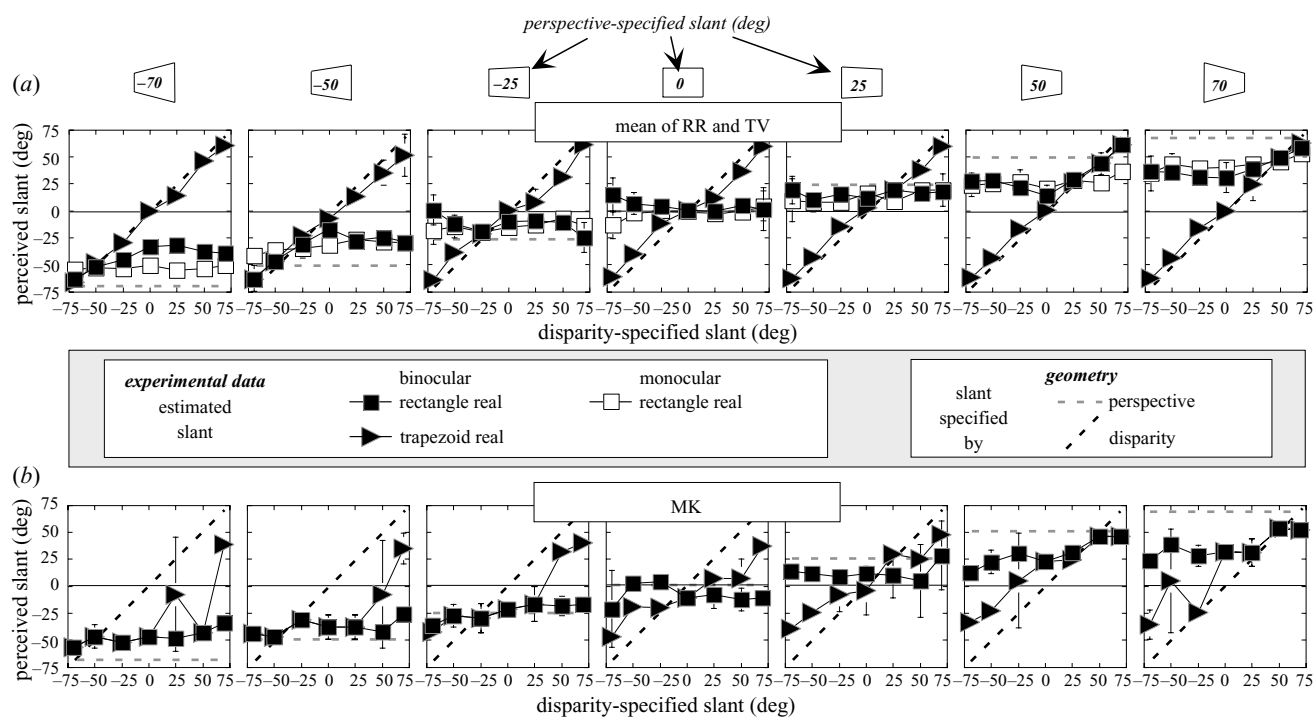
Figure 6. Data from experiment 1 for subjects who did not participate in the pictorial slant estimation of experiment 2. The symbols and the error bars denote the same as in figure 5. (a) Both the mean binocular and the mean monocular data of subjects RR and TV. Their data resembles the data of MS, NK and SV depicted in figure 5a. (b) MK's binocular data. Her slant estimations are hardly based upon disparity-specified slant.

reported two perceived slants. For this domain, subjects were able to voluntarily select one of the two perceived slants and to flip between them by switching their attention. It should be noted though that spontaneous flips could not be prevented, implying that the voluntary control was limited. The results show that in bi-stability, observers follow the disparity-predicted slant quite well when they perceive the trapezoid. When they perceive the rectangle, however, they do not follow the perspective-specified slant. In § 4 we will speculate on what these findings teach us. The plots for monocularly perceived slant (figure 5a(ii)) show, as expected, that the slant estimates do not vary with disparity. Generally, the data points are on a straight horizontal line parallel to the perspective-predicted slant. There are a couple of data points that deviate from the straight line, but their standard deviation is large. Note that the monocularly perceived slant was underestimated relative to the perspective-predicted slant.

For GK and SP (figure 5b) both the binocularly and the monocularly perceived slant estimates seem to be very similar to those of MS, NK and SV (figure 5a). For the binocular estimates there is, however, an interesting fundamental difference. GK and SP are able to perceive bi-stability even if disparity and perspective specify similar slants (see the data within the grey ellipses in figure 5). In § 4 we will relate the results of all subjects to their results in the stereo-anomaly test.

Figure 6a depicts the results of RR and TV. Except for two data points at perspective-specified slant −25° and 25°, their data are almost identical to those of MS, NK and SV (figure 5a) for both the binocular and monocular slant estimates. Figure 6b portrays the results of MK. For stimuli with huge differences between the disparity-specified and

the perspective-specified slant, she readily experienced perceptual bi-stability, but for the stimuli with smaller differences, she experienced bi-stability in only some trials. This was the case even with repeated presentation of the same stimulus, creating large standard deviations. When bi-stability could not be achieved, only one slant was observed, which was mainly based upon the perspective-specified slant. This pattern of data resembles the pattern of data in figures 5 and 6 for the perspective-dominated percept in bi-stability. In general, subject MK showed more difficulties in achieving bi-stability, which is probably related to her degraded stereo capacities.

(b) *Experiment 2*

In experiment 2, we examined the perceived pictorial slant created by the photographs of the slanted aluminium trapezoids of experiment 1. The grey diamond symbols in figure 5a depict the slant estimates of subjects MS, NK and SV and those in figure 5b the data of subjects GK and SP. The dark and light grey diamonds denote the binocularly and the monocularly estimated slants, respectively. For all participants, both binocular and monocular slant estimates are roughly on a horizontal line parallel to the perspective-specified slant. That is, they are independent of disparity. That the binocularly estimated slant is independent of disparity is not unexpected: linear perspective information in a photograph can only be exploited by making assumptions about the orientations of the object contours projected onto the image plane. As mentioned in § 2b(ii), the subjects were instructed that the photographs were made of slanted rectangles, but not of slanted trapezoids. The binocular estimates of GK and SP are generally larger than those of MS, NK and SV, both for binocular and monocular slant estimates. Figure 5a(ii) and b(ii) show

intriguing differences. For subjects MS, NK and SV the pictorial slant is generally smaller than the real rectangle slant, whereas the reverse is the case for subjects GK and SP.

4. DISCUSSION

We have examined the metrical aspects of voluntarily selected perceived slant in perceptual bi-stability for a broad spectrum of combinations of monocularly and binocularly specified slant. We have specifically studied perceived slant, both induced by real trapezoidal surfaces, and also in pictorial space of their photographed counterparts. The monocularly specified slants were signalled by perspective foreshortening. The binocularly specified slants were signalled by both disparity and perspective foreshortening. Most observers perceived only one slant when the monocularly and binocularly specified surface orientations were similar (in most cases meaning that the sign was identical). Observers were able to select either a monocularly or a binocularly dominated perceived slant when the specified orientations were rather different.

We found considerable behavioural differences between our subjects. For example, in the current study, two subjects (GK and SP) were able to experience mono stability only when disparity and perspective specified identical slants, whereas for the other subjects those slants needed to be only of the same sign to produce this experience (see figure 5). It is interesting to relate the differences between subjects to stereo-anomaly (Harwerth *et al.* 1998; van Ee & Richards 2002). Five of our subjects (NK, MS, RR, SV and TV) were excellent at distinguishing the signs and magnitudes of both the crossed and the uncrossed disparities. Two subjects (GK and SP) were able to correctly process crossed disparities (within a range of $-1°$ to $0°$), but not uncrossed disparities. This means that the reliability assigned to disparity-specified slant is probably smaller for GK and SP than for the other subjects. Thus, the contribution of the more interpretation-based, or more complex (Gillam & Cook 2001), perspective-specified slant relative to the disparity-specified slant is greater for GK and SP than for the other subjects. This, in turn, might mean that they are able to keep seeing the perspective interpretation, where other observers perceive only the (reconciled) disparity interpretation. Also possibly related to stereo-anomaly are the following findings. In our previous work on perceptual slant bi-stability with stereogram-produced slants, we generally found that, after the onset of the stimulus, observers first perceived the perspective-dominated slant (see Schriever (1925) for very similar findings in bi-stable slant from line drawings; and see van Ee *et al.* (2002) for a review on similar findings outside the domain of bi-stability). After a couple of seconds, the disparity-dominated percept 'kicked in'. Here, we found the same for subjects GK, MK and SP. However, MS, NK, RR, SV and TV first perceived the real slant of the trapezoid and, for them, it took a couple of seconds before the perspective slant was perceived. For our experiment with real planes it transpires that those subjects with excellent stereo acuities for both crossed and uncrossed disparities start seeing the disparity slant, the reverse being the case for subjects who are more perspective driven. There is another related interesting finding for pictorial slant. The slant estimates of GK and SP are generally larger than those of

MS, NK and SV, both for the binocular and for the monocular slant estimates. Further, figure 5a(ii), b(ii) shows that for MS, NK and SV the pictorial slant is generally smaller than the real rectangle slant, whereas the reverse is the case for subjects GK and SP. We speculated that this, too, might be consistent with the idea that GK and SP are more perspective-driven individuals than MS, NK and SV. In this view, the perceived pictorial slant is less hampered by conflicting disparity-specified slant for GK and SP than for the other subjects.

In two previous papers, data followed very similar patterns. In our first study on bi-stable slant perception (van Ee *et al.* 2002), bi-stability occurred only when the perspective- and the disparity-specified slants had opposite signs. In our second study (van Ee *et al.* 2003), and in our current study, bi-stability also occurred when the two cues had the same sign (but different amplitude). In the latter two studies, subjects were informed that the stimuli could be either trapezoidal or rectangular. In the first study, subjects were merely asked to report bi-stability. This difference in instructions could account for the slightly different position of the bifurcation from stable to bi-stable. Indeed, in the latter two studies observers commented that at one of the reported slants the object appeared trapezoidal and at the other reported slant the object appeared rectangular. In the latter two studies, subjects' disparity-dominated slant estimates were closer to the disparity-predicted slant than in the first study. In informal control experiments, we noticed that this is indeed related to the instruction. When naive subjects are asked to report whether they are able to perceive both a positive and a negative slant, they report smaller slants than when they are asked to report the slants of a trapezoid and a rectangle.

This brings us to the question: what do subjects mentally do when they attempt to substitute one percept for another voluntarily? One possible hypothesis is that observers are able to tap the binocularly perceived slant separately from the monocularly perceived slant. It is possible that two representations of the 3D layout coexist—a monocular representation and a binocular representation. One could even reason that the monocular representation of 3D space is a leftover from before the eyes migrated towards a frontal location. Usually in our daily vision we do not encounter situations in which the monocular and the binocular representation of the layout conflict, and therefore, we are not used to considering the two separate representations. However, in the laboratory, the two representations can be made apparent. Such a hypothesis could explain why, when subjects chose to perceive the trapezoid, slant estimates followed the disparity-predicted slant quite well, as if they chose a pure binocular representation of slant. For this binocular representation, all cues specify the same (real) slant once subjects relax the rectangularity assumption. When subjects chose to perceive the rectangle, however, the slant estimates are harder to explain by this hypothesis. Those estimates did not follow the monocularly perceived slant (compare the filled and the open squares in figure 5). In other words, it cannot be correct that observers are able to tap the binocularly perceived slant separately from the monocularly perceived slant. Apparently, subjects cannot turn off binocular vision at will and the estimated rectangle slant is a product of cue conflict in which disparity plays a part.

Concerning the rectangle, the perceived slant estimates were similar irrespective of whether slant was produced by the photograph or by the real trapezoid. The estimates are also the same as we obtained previously with stereogram-created slant (van Ee *et al.* 2003). The perspective interpretation requires a cognitive imagination of the slanted object that seems to be independent of how the retinal image is produced. What subjects do to interpret a perspective stimulus is something that we do all the time when we look at television or at pictures. When subjects estimate the rectangle slant (in binocular vision) they imagine, perhaps unconsciously, that they are looking at a picture of a slanted rectangular surface, in much the same way as we do when we look at the slanted door in figure 1*c*.

In summary, we have explored the metrical aspects of voluntarily selected perceived slant of real trapezoidal surfaces (for 'reverse-perspective' scenes) and their photographed counterparts (in pictorial space). Slant rivalry is not just a curiosity that occurs with flat stereograms (as the slant-contrast effect is). When subjects chose to perceive the trapezoid, the slant estimates followed the disparity-predicted slant with only a slight underestimation, as if they chose a pure binocular representation of slant. When subjects chose to perceive the rectangle, their estimates were similar irrespective of the way the retinal images had been created (by real or by photographed surface outlines).

REFERENCES

Ames, A. 1951 Visual perception and the rotating trapezoidal window. *Psychol. Monogr. Gen. Appl.* **65**, 1–31.

Blake, R. & Logothetis, N. K. 2002 Visual competition. *Nature Rev. Neurosci.* **3**, 1–11.

Clark, W. C., Smith, A. H. & Rabe, A. 1956 The interaction of surface texture, outline gradient, and ground in the perception of slant. *Can. J. Psychol.* **10**, 1–8.

Cook, N. D., Hayashi, T., Amemiya, T., Suzuki, K. & Leumann, L. 2002 Effects of visual-field inversions on the reverse-perspective illusion. *Perception* **31**, 1147–1151.

Ellis, S. R., Kaiser, M. K. & Grunwald, A. C.(eds) 1991 Pictorial communication in virtual and real environments. London: Taylor & Francis.

Frisby, J. P., Buckley, D. & Horsman, J. M. 1995 Integration of stereo, texture, and outline cues during pinhole viewing of real ridge-shaped objects and stereograms of ridges. *Perception* **24**, 181–198.

Gillam, B. J. 1968 Perception of slant when perspective and stereopsis conflict: experiments with aniseikonic lenses. *J. Exp. Psychol.* **78**, 299–305.

Gillam, B. J. & Cook, M. L. 2001 Perspective based on stereopsis and occlusion. *Psychol. Sci.* **12**, 424–429.

Harwerth, R. S., Möller, M. C. & Wensveen, J. M. 1998 Effects of cue context on the perception of depth from combined disparity and perspective cues. *Optom. Vision Sci.* **75**, 433–444.

Hecht, H., van Doorn, A. J. & Koenderink, J. J. 1999 Compression of visual space in natural scenes and in their photographic counterparts. *Percept. Psychophys.* **61**, 1269–1286.

Howard, I. P. & Rogers, B. J. 2002 *Seeing in depth vol 2: depth perception*. Toronto: I. Porteous.

Koenderink, J. J. & Van Doorn, A. J. 2003 Pictorial space. In *Looking into pictures: an interdisciplinary approach to pictorial space* (ed. H. Hecht, R. Schwartz & M. Atherton, pp. 239–403.) Cambridge, MA: MIT Press.

Kubovy, M. 1986 *The psychology of perspective and Renaissance art*. Cambridge University Press.

Mach, E. 1866 Über die physiologische Wirkung räumlich verteilter Lichtreize. *Sitzungsb. Wiener Akad.* **54**, 3.

Papathomas, T. V. 2000 See how they turn: false depth and motion in Hughes' reverspectives. *Hum. Vision Electr. Imag. V, SPIE Proc. Ser.* **3959**, 506–517.

Papathomas, T. V. 2002 Experiments on the role of painted cues in Hughes's reverspectives. *Perception* **31**, 521–530.

Reinhardt-Rutland, A. H. 1990 Detecting orientation of a surface: the rectangularity postulate and primary depth cues. *J. Gen. Psychol.* **117**, 391–401.

Rosinski, R. R., Mulholland, T., Degelman, D. & Farber, J. 1980 Picture perception: an analysis of visual compensation. *Percept. Psychophys.* **28**, 521–526.

Schriever, W. 1925 Experimentelle Studien über stereoskopisches Sehen. *Zeitschr. Psycholo. Physiol. Sinnesorgane* **96**, 113–170.

Slyce, J. 1998 *Patrick Hughes: perverspective*. London: Momentum.

van Doorn, A. J., Koenderink, J. J. & de Ridder, H. 2001 Pictorial space correspondence in photographs of an object in different poses. In *Human Vision and Electronic Imaging VI: Proc. Soc. Photo-Optical Instr. Eng.*, vol. 4299 (ed. B. E. Rogowitz & T. N. Pappas) pp. 321–329. Washington, DC: SPIE.

van Ee, R. 2005 Dynamics of perceptual bi-stability for stereoscopic slant rivalry and a comparison with grating, house-face, and Necker cube rivalry. *Vision Res.* **45**, 29–40.

van Ee, R. & Richards, W. 2002 A planar and a volumetric test for stereoanomaly. *Perception* **31**, 51–64.

van Ee, R., Banks, M. S. & Backus, B. T. 1999 An analysis of binocular slant contrast. *Perception* **28**, 1121–1145.

van Ee, R., van Dam, L. C. J. & Erkelens, C. J. 2002 Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. *J. Vision* **2**, 597–607.

van Ee, R., Adams, W. J. & Mamassian, P. 2003 Bayesian modelling of perceived slant in bi-stable stereoscopic perception. *J. Optic. Soc. Am.* **20**, 1398–1406.

van Ee, R., van Dam, L. C. J. & Brouwer, G. J. 2005 Voluntary control and the dynamics of perceptual bi-stability. *Vision Res.* **45**, 41–55.

Wade, N. J. & Hughes, P. 1999 Fooling the eyes: *trompe l'oeil* and reverse perspective. *Perception* **28**, 1115–1119.

Werner, H. 1937 Dynamical theory of depth perception. *Psychol. Monogr.* **49**, 1–127.

Yellott, J. I. & Kaiwi, J. L. 1979 Depth inversion despite stereopsis: the appearance of random-dot stereograms on surfaces seen in reverse perspective. *Perception* **8**, 135–142.