Shape-from-shading depends on visual, gravitational, and body-orientation cues

Heather L Jenkin, Michael R Jenkin, Richard T Dyde, Laurence R Harris¶
Centre for Vision Research, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada;
e-mail: harris@yorku.ca
Received 1 February 2004, in revised form 18 March 2004; published online 4 November 2004

Abstract. The perception of shading-defined form results from an interaction between shading cues and the frames of reference within which those cues are interpreted. In the absence of a clear source of illumination, the definition of 'up' becomes critical to deducing the perceived shape from a particular pattern of shading. In our experiments, twelve subjects adjusted the orientation of a planar disc painted with a linear luminance gradient from one side to the other, until the disc appeared maximally convex—that is, until the luminance gradient induced the maximum perception of a three-dimensional shape. The vision, gravity, and body-orientation cues were altered relative to each other. Visual cues were manipulated by the York Tilted Room facility, and body cues were altered by simply lying on one side. The orientation of the disc that appeared maximally convex varied in a systematic fashion with these manipulations. We present a model in which the direction of perceptual 'up' is determined from the sum of three weighted vectors corresponding to the vision, gravity, and body-orientation cues. The model predicts the perceived direction of 'up', contributes to our understanding of how shape-from-shading is deduced, and also predicts the confidence with which the 'up' direction is perceived.

1 Introduction

In the absence of a clear source of illumination, light is assumed to come from 'above' (Mamassian and Goutcher 2001; Ramachandran 1988). A corollary of this is that those shading gradients over a surface in the 'above – below' direction will make the most significant contribution to the perceived three-dimensional (3-D) shape of that surface. Identical gradients that occur orthogonal to the 'above – below' direction and which do not correspond to an assumed illumination source, are not interpreted as shading and do not contribute to the perceived surface shape. The perception of 3-D shape-from-shading therefore depends not only on a luminance gradient but also upon the perceived direction of 'up'. In fact, these cues may be more important than actual illumination direction, even when the lighting source is clear (Mingolla and Todd 1986).

The influence of various perceptual frames on the perception of 'up' has been investigated both for humans (eg Howard et al 1990; Yonas et al 1979) and for animals (eg Hershberger 1970). Yonas et al (1979) reported that 4-year-old infants use the head more than gravity as the frame of reference in interpreting surface relief, but that children aged seven years make about equal use of the two frames of reference. For adults, the assumption about the direction of illumination is predominantly with respect to the head (Howard et al 1990; Kleffner and Ramachandran 1992). But the role of head may have been overstated (eg by Connor 2001): perceived vertical is also influenced by the orientation of the body (Mittelstaedt 1983), the physical direction of gravity, and by visual cues (Jenkin et al 2002, 2003a; Oman 2003). In other words, 'up' is neither fixed nor objective and can be manipulated (sometimes unintentionally) by altering the relationship of the visual, gravity, and body-defined frames.

A mismatch of visual environmental cues and gravity can, for example, occur when watching TV or a film where the camera is tilted or rolled. Even looking at a picture in which the local horizon does not match the actual horizon can provide

¶ Author to whom all correspondence should be addressed.

visual cues to 'up' that are mismatched with body and gravity cues. Lying down can also generate such a mismatch. If maneuvres as simple as looking at a tilted picture or lying down can alter the perceived direction of 'up', then they can also alter the expected direction of shading and potentially influence perceptions that derive from this, especially shape-from-shading.

If the perception of shape-from-shading is tied to the direction of 'up', then changing the perceived 'up' direction should alter perceived shape. In this paper, we alter the perceived 'up' direction by separating visual, gravity, and body-defined frames and examine the effect of these manipulations on the perception of shape-from-shading. The body frame can be further divided into torso, head, and retinally defined frames but these have been kept aligned in the present experiments and their individual contributions cannot therefore be distinguished. We have previously reported the influence of frame manipulation on the perception of the 'up' direction using a constant stimuli paradigm (Harris et al 2002; Jenkin et al 2003a, 2003b). Here we use an adjustment task to explore the contribution of these frames to the perception of shape-from-shading directly.

2 Methods
2.1 Overview
Twelve subjects (nine male, three female, aged 22 – 50 years) were asked to adjust the roll orientation of a disc that varied in luminance from one side to the other, until it appeared maximally convex. The disc was presented on the screen of a laptop computer positioned orthogonally to the subject's line of sight. Visual orientation cues were manipulated by viewing the screen either in a normal room or inside the York Tilted Room facility (figure 1), a room constructed at 90° to the normal orientation. Body-orientation cues were manipulated by conducting the judgments with subjects either sitting upright or lying down on one side.
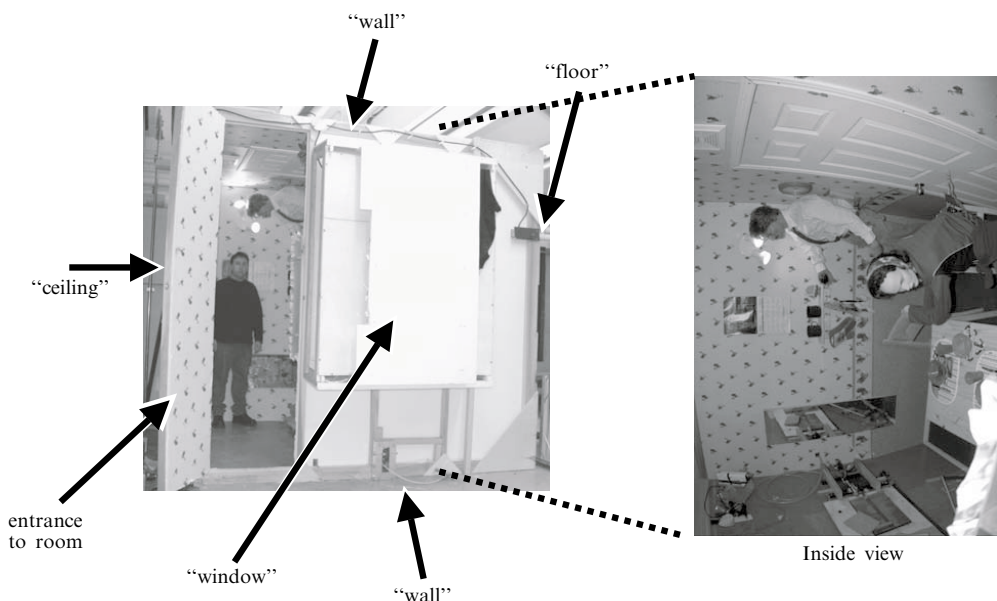

Figure 1. The York Tilted Room. The room is a 2.4 m × 2.4 m × 2.4 m room tilted at 90° to gravity. The visual contents of the room are highly polarised to provide strong intrinsic cues about a visually defined 'up' direction that is orthogonal to gravity. The photograph on the left shows the room from the outside with the door open and an experimenter standing inside to show the actual orientation and scale. Note the head of the mannequin visible through the door. The door is closed during experiments. The photograph on the right shows the inside view.

2.2 *The York Tilted Room*

The York Tilted Room is 2.4 m on each side and decorated like a real room with many objects placed in their natural intrinsic relationships with each other, except that everything is arranged to indicate 'down' at 90° to the normal orientation. The wallpaper has a strongly polarised pattern, there are books on the bookshelves, nicknacks on the windowsill, and place settings on the table. The room has been constructed in a tilted orientation, so that the visual floor is one of the physical walls, and one of the visual walls appears on the physical ceiling (figure 1).

2.3 *Procedure*

Observers either sat upright, or lay on their right sides, either in the tilted room or in a normally oriented room (figure 2). For each of these four conditions, participants were shown a single shaded disc on a computer screen that was masked to a square that subtended 28 deg at the viewing distance of 30 cm, with the room clearly visible behind the computer display. The disc had an initial orientation of the shading gradient axis set pseudo-randomly to one of 120 possible positions evenly spaced around the clock face. For each of the 120 trials the subjects' task was to rotate the disc by pressing buttons on the game pad (which rotated the disc in 1° steps) until the perceived convexity of the disc was greatest. When the subjects had reached the decision, they pressed a third button on the game pad, whereupon the software recorded this orientation and initiated the next trial. Subjects controlled the length of each trial but moved through the trials quickly, finishing the 120 trials in each condition in less than 10 minutes.
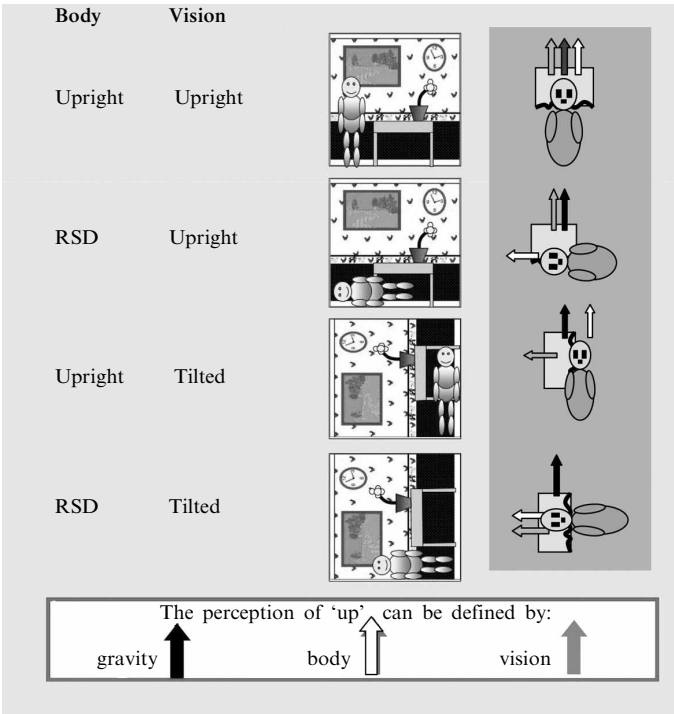

Figure 2. The four experimental conditions. Subjects were tested upright in an upright room (top row), lying right side down (RSD) in an upright room (second row), upright in a tilted room (third row), and lying right side down (RSD) in a tilted room (bottom row). These arrangements altered the relative directions of gravity, body, and visual orientation cues. The filled arrows represent the direction of gravity, the open arrows represent the orientation of the body, and the gray arrows represent the orientation of vision.

3 Results

The frequencies with which each orientation of the disc was chosen as having the maximum perceived convexity for each condition were normally distributed (figure 3a). The best-fit Gaussians for the frequency distributions are plotted through the data in both linear (figure 3a) and polar (figure 3b) coordinates, and the values for the peaks and standard deviations are given in table 1. A two-way repeated-measures ANOVA of the mode of each subject's response showed a significant effect of whether the room was tilted or not ($F_{1,11} = 9.57$, $p < 0.01$) and whether the subject was lying right side down or upright ($F_{1,11} = 16.56$, $p < 0.01$). There were no significant interactions between these two ($F_{1,11} = 0.11$, ns), which supports a linear model of the effects of these influences on the interpretation of shading produced by tilting either the room or the person relative to gravity.
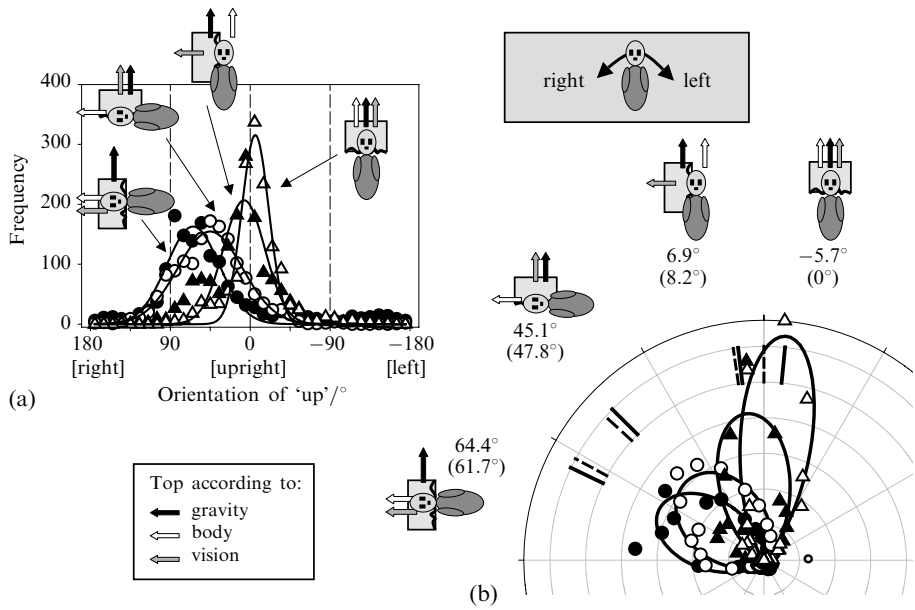

Figure 3. Frequency distributions of the number of times each orientation of the test disc was chosen as evoking the impression of being the most convex in 10° bins. (a) Linear plot with the *x*-axis representing the orientation of the test disc relative to gravity. 0° corresponds to upright, positive numbers correspond to a right tilt of the chosen orientation relative to gravity. The four experimental conditions are shown as inserts (conventions as in figure 2). Open triangles: upright in an upright room; closed triangles: upright in the tilted room; open circles: lying right side down in an upright room; and closed circles: lying right side down in the tilted room. Best-fit Gaussians are plotted as solid lines through the data. (b) The same data and the best-fit Gaussians plotted in polar coordinates (relative to gravity), with counterclockwise rotations corresponding to roll to the right. The numbers specify the peaks of the four Gaussians which are also indicated by the solid lines radiating out at the edge of the polar plot. The numbers in parentheses and the dashed radiating lines indicate the values from the weighted linear summation model (see text).

4 Discussion

These experiments have confirmed previous findings (eg Ramachandran 1988) that shading affects perceived shape: a disc with a luminance gradient across it appears 3-D and the extent of the perceived convexity or concavity depends on the direction of illumination. Our results extend this finding and indicate that the apparent convexity of the disc appears greatest when the gradient is aligned with the perceived up/down direction. This alignment appears essential for the luminance gradient to be interpreted

Table 1. The orientation of a shaded disc that appeared most convex under our experimental conditions and the standard deviation of the 1440 settings made by our twelve subjects for each condition. The direction and length of the output of the linear summation model are given in the right two columns. All values are reported relative to the direction of physical gravity (see figure 3b).

Condition	Orientation of maximum perceived convexity[a]	Standard deviation of subjects' settings	Linear model	
			direction of vector	length of vector
body upright, upright room	5.7° left	15.3°	0°	2.86
right side down, upright room	45.1° right	36.4°	47.8° right	2.02
body upright, tilted room	6.9° right	23.5°	8.2° right	2.52
right side down, tilted room	64.4° right	29.7°	61.7° right	2.11

[a] 0° aligned with gravity.

as shading and to create the impression of a 3-D form. This might explain why varying the illumination direction while keeping cues to 'up' constant is relatively ineffective in altering perceived object shape (Vogels and Biederman 2002). When a circle with a luminance gradient was rotated away from the perceived up/down orientation, the effectiveness of the gradient as a cue to shape was rapidly reduced. When the luminance gradient was orthogonal to the perceived up/down direction, the disc appeared flat. The perceived up/down direction depended on the orientation of vision, gravity, and body cues and was not dominated by any single cue.

4.1 Predicting the direction of luminance gradient interpreted as shading

A simple model of how visual, gravity, and body-orientation cues can be combined to generate the perceived direction of 'up', is to take the weighted vector sum of their directions. This is shown graphically in figure 4 and in the following equation:

$$\mathbf{up} = \mathbf{k}_v + \mathbf{k}_g + \mathbf{k}_b \,, \tag{1}$$

where $\mathbf{k}_v$, $\mathbf{k}_g$, and $\mathbf{k}_b$ are the relative lengths (or weightings) of the vectors corresponding to vision, gravity, and body orientations, respectively. Note that under this model, light comes from directly above when upright in a visually upright environment.

The direction of the **up** vector (θ) relative to the direction of physical gravity can therefore be specified for our conditions (as shown in figure 4) as:

$$\theta = \arctan\left(\frac{k_b + k_v}{k_g}\right) \text{ for the right-side-down, tilted-room condition;} \tag{2}$$

$$\theta = \arctan\left(\frac{k_b}{k_v + k_g}\right) \text{ for the right-side-down, upright-room condition;} \tag{3}$$

$$\theta = \arctan\left(\frac{k_v}{k_b + k_g}\right) \text{ for the upright, tilted-room condition .} \tag{4}$$

Note that being upright in an upright-room condition does not provide any constraints on $\mathbf{k}_v$, $\mathbf{k}_g$, and $\mathbf{k}_b$ as all the vectors are aligned. Since the lengths of the vision, gravity, and body vectors are specified relative to each other, we can arbitrarily set one of them ($\mathbf{k}_g$) to unity in each equation, and obtain a linear constraint on the remaining values, $\mathbf{k}_v$ and $\mathbf{k}_b$. The three resulting constraints are plotted in figure 5. The least-squares solution
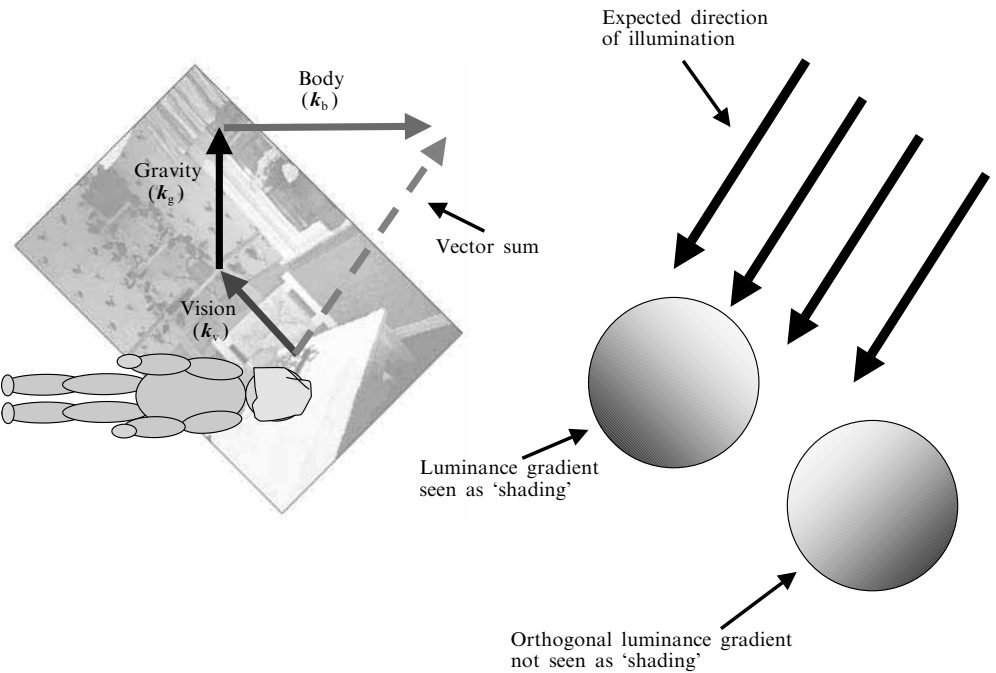
Figure 4. A linear summation model for determining the perceived direction of 'up' from a weighted sum of three vectors representing vision (k_v), the body (k_b), and gravity (k_g). The vector sum of these three indicates the direction of 'up' (dashed line). The perceived direction of 'up' corresponds to the direction from which illumination is expected (array of solid arrows parallel to the dashed line). The model thus predicts the expected orientation of the luminance gradient due to shading under the conditions shown on the left of the figure, to be as shown on the upper disc. For this combination of gravity, body, and visual orientation, the luminance gradient on the lower disc, which is orthogonal to the expected direction, would not be interpreted as shading and would not contribute to the perceived shape of the surface.
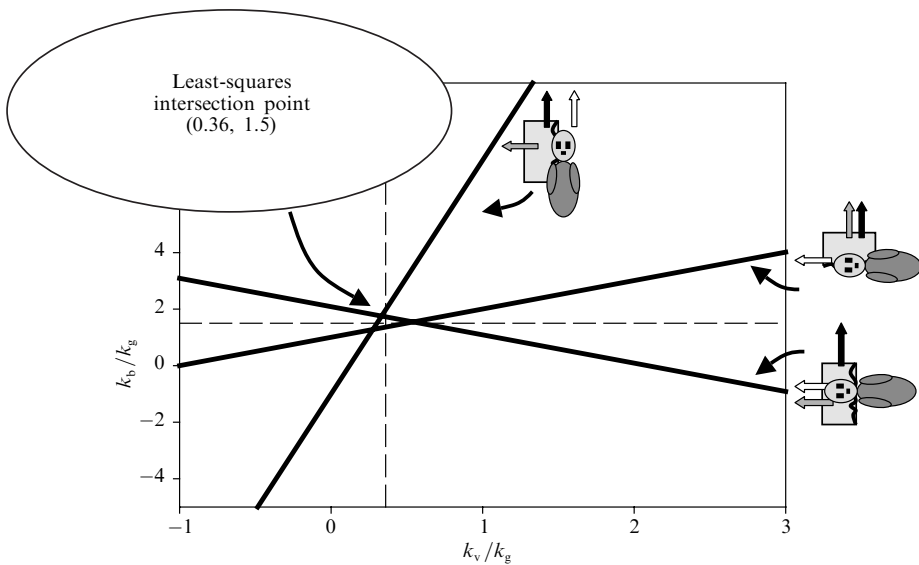

Figure 5. The linear constraints on the ratios k_v/k_g and k_b/k_g from the settings made under each of the three conditions shown in the cartoons (conventions as for figure 2). The lines almost intersect at a single point, indicating very little variability between conditions in estimating the values of k_v and k_b.

to the three curves (the point that best approximates where they intersect) is given by:

$$\text{vision : body : gravity} = \boldsymbol{k}_{\text{v}} : \boldsymbol{k}_{\text{b}} : \boldsymbol{k}_{\text{g}} = 0.36 : 1.5 : 1 \,.$$

These values are similar to those reported by us previously (Jenkin et al 2003a), although the relative weighting of vision reported here is lower. In the previous experiment, we extrapolated the orientation most likely to correspond to maximum curvature from the probability with which each of four discs with fixed orthogonal gradients was chosen as 'most convex'. The weightings of the contributing factors to the perception of 'up' can be used to predict the direction of 'up' for any combination of visual, body, and gravity orientations (table 1 and figure 3), and suggest that a simple weighted-sum-direction model is an adequate model of how the brain estimates the 'up' direction.

The length of the vector defined in equation (1) specifies the strength of the proposed internal signal corresponding to perceptual representation of 'up'. Figure 6 shows the standard deviation of the settings obtained under each of our four conditions against the length of this vector. The correlation coefficient is 0.95 indicating that the linear summation model is a very good predictor not only of the direction but also the variability of the perceived 'up' direction.
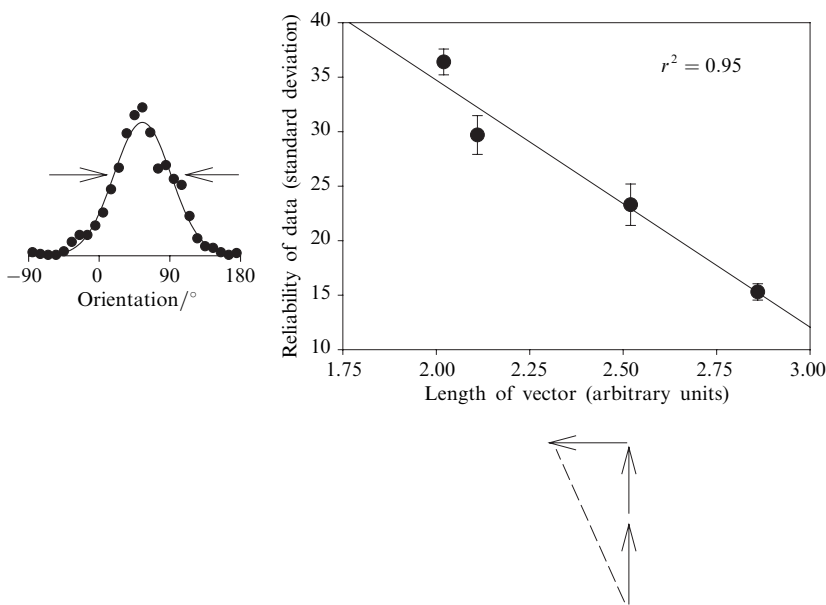

Figure 6. Correlation between the length of the vector of the linear summation model (dashed line, see figure 4 and text) and the uncertainty of the subjects' setting of the 'up' direction, as given by the standard deviation of their settings (see insert and figure 3). Error bars are the standard errors of the parameter fit. The regression coefficient (r^2) is 0.95 suggesting that the length of the vector sum determines the confidence with which the 'up' direction is perceived.

The linear summation model does not explain the small (5.7°) shift to the left found when subjects were upright in an upright room, since it is bound to predict 'up' as being in the direction of the vectors when they are all aligned. Mamassian and Goutcher (2001) also reported a leftward bias with upright subjects, supported by neurophysiological correlates (Mamassian et al 2003). This bias may reflect a tilt in the internal representation in one of the vectors and might be connected to a general attentional bias towards the right side of space (Spence et al 2001).

5 Conclusion

As shape-from-shading is intimately related to the perceived 'up' direction, and as we can manipulate the perceived 'up' direction through appropriate manipulations of the visual, body, and gravity frames, we can manipulate, in a controlled manner, the perceived 3-D shape of a stimulus with a fixed luminance gradation across it. Although the experiment reported above utilised a full room to generate the visual display, viewing a photograph with strong polarisation cues indicating 'up' and 'down' can also, perhaps surprisingly, provide cues to visual orientation that are sufficient to alter the perceived direction of 'up'. A simple demonstration of this is shown in figure 7. The figure shows a photograph with very clear cues to orientation on which are super-imposed four shaded discs. (Our experiments used only a single disc but the effect can be more dramatically demonstrated when there are four in the complementary arrangement shown.) By viewing figure 7, first upright and then with the body (at least the head) on one side, it can be quickly seen that the orientation of the page (relative to either gravity or the head) that produces the strongest appearance of 3-D convexity depends on the orientation of the head relative to gravity.

Figure 7. The orientation at which the discs in the figure appear maximally convex and concave can be predicted from the instantaneous vector sum of body, gravity, and the orientation of the background photograph. When this direction is aligned with the orientation of the shading gradient, the circles will appear maximally convex and concave. To demonstrate this, adopt an upright pose and adjust the orientation of the printed page to generate the greatest convexity and concavity in the four central discs. Note the orientation you have chosen relative to gravity (and your head). Then repeat the adjustment with your head tilted to one side and compare the setting. The difference in these orientations illustrates the importance of the relative directions of these three cues to the perception of 'up' and the consequent effect of this on the perception of shape-from-shading.

Our results show that the contribution to the perception of form derived from shading depends on the perceived direction of 'up'. That component of a luminance gradient which aligns with this direction is most effective in determining perceived shape.

References
Connor C E, 2001 "Visual perception: sunny side up" *Current Biology* **11** R776 – R778
Harris L R, Jenkin H L, Dyde R T, Kaiserman J, Jenkin M R, 2002 "Visual and vestibular cues in judging the direction of 'up'" *Journal of Vestibular Research* **11** 307 (BP12.4)
Hershberger W, 1970 "Attached-shadow orientation perceived as depth by chickens reared in an environment illuminated from below" *Journal of Comparative Physiology and Psychology* **73** 407 – 411
Howard I P, Bergström S S, Ohmi M, 1990 "Shape from shading in different frames of reference" *Perception* **19** 523 – 530
Jenkin H L, Dyde R T, Jenkin M R, Harris L R, 2002 "Judging the direction of 'above' in a tilted room" *Perception* **31** Supplement, 30 (abstract)
Jenkin H L, Dyde R T, Jenkin M R, Harris L R, Howard I P, 2003a "Relative role of visual and non-visual cues in judging the direction of 'up': Experiments in the York Tilted Room facility" *Journal of Vestibular Research* **13** 287 – 293
Jenkin H L, Dyde R T, Zacher J E, Jenkin M R, Harris L R, 2003b "Multi-sensory contributions to the perception of up: Evidence from illumination judgements" *Journal of Vision* **3** 638a
Kleffner D A, Ramachandran V S, 1992 "On the perception of shape from shading" *Perception & Psychophysics* **52** 18 – 36
Mamassian P, Goutcher R, 2001 "Prior knowledge on the illumination position" *Cognition* **81** B1 – B9
Mamassian P, Jentzsch I, Bacon B A, Schweinberger S R, 2003 "Neural correlates of shape from shading" *NeuroReport* **14** 971 – 975
Mingolla E, Todd J T, 1986 "Perception of solid shape from shading" *Biological Cybernetics* **53** 137 – 151
Mittelstaedt H, 1983 "A new solution to the problem of the subjective vertical" *Naturwissenschaften* **70** 272 – 281
Oman C M, 2003 "Human visual orientation in weightlessness", in *Levels of Perception* Eds L R Harris, M R Jenkin (New York: Springer) pp 375 – 398
Ramachandran V S, 1988 "The perception of shape from shading" *Nature* **331** 163 – 166
Spence C, Shore D I, Klein R M, 2001 "Multisensory prior entry" *Journal of Experimental Psychology: General* **130** 799 – 832
Vogels R, Biederman I, 2002 "Effects of illumination intensity and direction on object coding in macaque inferior temporal cortex" *Cerebral Cortex* **12** 756 – 766
Yonas A, Kuskowski M, Sternfels S, 1979 "The role of frames of reference in the development of responsiveness to shading information" *Child Development* **50** 495 – 500

p

PERCEPTION
VOLUME 33 2004