

# Neural correlates of shape from shading

Pascal Mamassian,<sup>CA</sup> Ines Jentzsch, Benoit A. Bacon and Stefan R. Schweinberger

Department of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, UK

<sup>CA</sup>Corresponding Author: pascal@psy.gla.ac.uk

Received 26 January 2003; accepted 26 February 2003

DOI: 10.1097/01.wnr.0000069061.85441.f2

Psychophysical studies have shown that human observers resolve shape-from-shading ambiguities by assuming that light is coming from above-left. Using event-related potentials (ERPs), we measured the processing time of the perception of an ambiguous shaded pattern. We found that the N2 component followed the change of perceived shape with stimulus orientation. We also found that the P1 component in occipital and temporal areas was corre-

lated with the observers' idiosyncratic bias for light source position. The precocity of the correlated ERP components suggests that the light source is represented early in the visual system. Altogether, our results indicate that shape-from-shading is a mostly bottom-up mechanism. *NeuroReport* 14:971-975 © 2003 Lippincott Williams & Wilkins.

**Key words:** Ambiguous perception; Brain potential correlates; Depth perception; Event-related potentials; Monocular vision; Prior knowledge; Shape-from-shading

## INTRODUCTION

The visual system relies on multiple cues to infer the 3D structure of the environment. In addition to binocular and motion cues, pictorial cues also play a major role in depth perception when the organism is static and looking at the world monocularly. Recent studies have started to identify the cortical areas involved in the processing of depth from these pictorial cues. For instance, Taira and colleagues [1] used fMRI to isolate an area in the right intraparietal cortex that seems to be involved in the discrimination of 3D shape-from-shading. This area in humans is believed to be the homologue of area CIP in the lateral bank of the intraparietal sulcus (IPS) in monkeys where cells respond to both linear perspective and binocular disparities [2] and have similar tuning for slanted planes defined by texture gradients and binocular disparities [3]. Since these cells code for 3D structure irrespective of the depth cue providing that information, it is logical to infer that these cues are first processed in lower visual areas prior to their integration in IPS. There is indeed evidence from two studies that early visual cortical areas such as V1 and V2 are involved in depth perception from pictorial cues [4,5]. Both studies used similar shape-from-shading stimuli that evoked a pop-out percept of convex hemispheres amongst concave ones [6]. However, it is not clear whether the selectivity of neurons in these early visual areas results from bottom-up or top-down processing. While bottom-up processing is usually believed to be mostly autonomous, top-down processing reflects feedback of higher cortical areas onto earlier ones and depends on attentional and other task-specific factors. Since pop-out stimuli imply attentional mechanisms, it is difficult to judge from these previous

studies [5] whether shape-from-shading arises from bottom-up or top-down processing. The purpose of the present study was to reveal the importance of bottom-up processing for shape-from-shading with a more direct shape discrimination task.

The perception of shape-from-shading is ambiguous because multiple 3D shapes can produce the same shading pattern. For instance, a concave shape lit from below can display the same shading pattern as a convex shape lit from above. Some of these ambiguities are removed when the position of the light source is known. It is now well established that humans assume by default that light is located above their head [6]. Moreover, more recent studies have revealed a bias to the left for the assumed light source position, meaning that the preferred light position is above-left rather than straight above [7,8]. The cause of this leftward bias remains unknown.

How is the default assumption about light source position encoded by the visual system? Intuitively, it may seem that light source position is a fairly cognitive concept that could play a role in top-down processing. For example, if I hypothesise that light comes from the left, then I can interpret some black regions in the image as shadows and thereby infer the shape of the objects casting the shadows. Alternatively, it may be argued that any visual assumption can really be advantageous if it is available early on and therefore can be of use in bottom-up processing. The advantage of such an early assumption is that the resulting interpretation would be obtained quickly and could eventually be re-evaluated in face of further evidence.

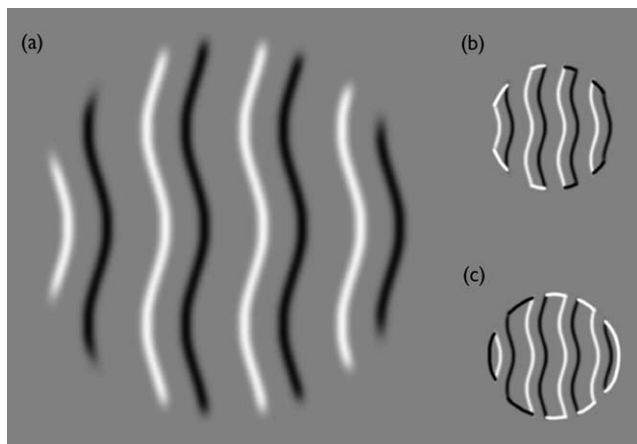
We investigated how early the assumption about light source was encoded by measuring the time course of

visually evoked brain activity in healthy human observers. We report here data from visual evoked potentials recorded on the scalp of observers while they were presented with ambiguous shaded patterns.

## MATERIALS AND METHODS

**Participants:** Fourteen participants were recruited from the research community of the Psychology Department at the University of Glasgow. While all observers were able to see the 3D shape of a shaded pattern when it was presented for a long time in the fovea, four observers were discarded for failing to see depth in the stimulus when it was flashed briefly in the periphery. Of the remaining ten participants, one displayed large horizontal eye-movements in the direction of the stimulus that disrupted the ERPs and another showed a bias to the left for the light position that was  $> 3$  s.d. from the mean of biases. All the analyses presented below were therefore carried out on the remaining eight participants. Of these eight observers, four were the authors while the other four were naive as to the purpose of the experiment. Their mean age was 29.5 years, ranging between 22 and 38. All observers had normal or corrected to normal eyesight.

**Stimuli:** The stimuli were similar to the ones used by Mamassian and Goutcher [8] except that they were smaller (diameter  $2.0^\circ$  of visual angle) to enable us to present them at different eccentricities. They represented a fronto-parallel surface with strips in relief that appeared either narrow or wide depending on the assumed light position (Fig. 1). Stimuli were presented for 200 ms in order to avoid eye movement to eccentric positions. Stimuli were presented either centrally or to the left or right of the fixation point ( $3.0^\circ$  of eccentricity) and at 16 orientations in the image plane, completely randomized within blocks. Each orientation



**Fig. 1.** Example of stimuli. The stimuli consisted of alternating black and white undulating lines on a grey background. Black lines were usually seen as shadow edges while white lines were seen as edges facing the light source. This stimulus was presented at 16 orientations in random order. In the orientation shown in (a), the stimulus can be interpreted as an object with narrow strips in relief if light is assumed to come from the left (b), or as an object with wide strips in relief if light comes from the right (c).

at each eccentricity was presented a total of 32 times. To counterbalance a potential effect of the lines closest to the fixation point, half of the stimuli had their centre on a wide bar while the other half had its centre on a narrow bar. Because this control did not produce any noticeable difference, these two sets of stimuli are combined in the analysis presented below. In summary, the stimuli varied in orientation (16 levels) and eccentricity (left, centre, right).

**Apparatus:** The stimuli were presented on a 17-inch monitor placed 1 m from the observer. Observers placed their head on a chin-rest to minimise head movements. Electrophysiological recordings were made using Neuroscan Synamps DC amplifiers.

**Procedure:** The experiment was conducted in accordance with the Declaration of Helsinki and with the agreement of the local ethics committee. Participants gave their prior consent and had the option to quit the experiment at any time.

The participants were asked to fixate the fixation point and to look at the stimuli without blinking or moving their eyes. Their task was to decide whether the stimulus presented contained narrow or wide strips in relief (two-alternative forced-choice procedure). In order to minimise the contamination of ERP signals by the motor response, the participants were asked to withhold their response until the stimulus and the subsequent mask disappeared.

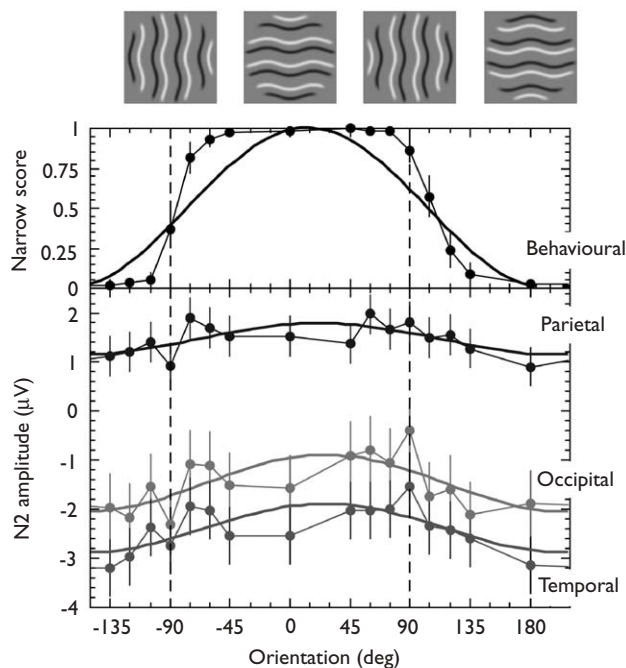
**Electrophysiological recordings:** Electroencephalographic (EEG) activity was continuously recorded with sintered Ag/AgCl electrodes mounted in an electrode cap (Easy-Cap) at the scalp positions Fz, Cz, Pz, Iz, FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, FT9, FT10, P9, P10, PO9, PO10, C3', C4', F9', F10' and TP9. The right mastoid (TP10) served as initial common reference, and the AFz electrode as ground. The C3' electrode was positioned 0.75 cm anterior to the midpoint of a straight line between C3 and C1, and the C4' electrode between C4 and C2. The F9' and F10' electrodes were positioned 2 cm anterior to F9 and F10 at the outer canthi of the left and right eye. Vertical electroocular (vEOG) was bipolarly registered above and below the right eye. EEG and EOG recordings were sampled at 250 Hz. Electrode impedance was kept below 10 k $\Omega$ . All signals were recorded in DC mode, with low-pass filters set to 40 Hz ( $-6$  dB attenuation, 12 dB/octave).

Off-line analysis epochs were generated, starting 200 ms prior to stimulus onset and lasting for a total duration of 1000 ms. Epochs were aligned to a 200 ms pre-stimulus baseline. Automatic artifact detection software was run for an initial sorting of trials and all trials were then visually inspected for artifacts of ocular (e.g. blinks, saccades) and non-ocular origin (drifts, channel blockings). Blink trials were corrected and all trials with saccades or non-ocular artifacts were discarded. The epochs were averaged separately for each channel and each experimental condition. The resulting ERPs were low-pass filtered at 10 Hz and re-referenced to average reference, excluding the vEOG channel.

## RESULTS

**Behavioural results:** We computed the narrow score, that is the proportion of times that an observer indicated seeing narrow strips in relief for each stimulus condition. Narrow score values close to one and zero corresponded to stable percepts of narrow and wide strips, respectively, whereas narrow scores close to 0.5 corresponded to bi-stable percepts. As expected from our previous study where stimulus presentations were only made centrally [8], narrow scores were significantly affected by the orientation of the stimulus for all three eccentricities ( $F(15,105) > 30.9$ ,  $p < 0.05$ ).

Maxima of the narrow scores were obtained for stimulus orientations close to zero. At such orientations, the stimulus is consistent with narrow strips in relief only if light is coming from above, which indicates that observers assumed light to come from above to interpret the stimuli. Similarly, minima of the narrow scores were obtained for stimulus orientations close to 180 degrees, and at these orientations the stimulus is consistent with wide strips in relief only if light is coming from above. Examining more closely the variation of narrow scores with stimulus orientation, we notice that the extrema of the narrow scores are in fact shifted slightly to the right (Fig. 2, top graph). This shift is consistent with an assumed light position located above-left rather than straight above. The magnitude of this leftward bias can be estimated by fitting the data with a raised



**Fig. 2.** Behavioural and physiological results. The behavioural results (top graph) show the effect of stimulus orientation on the narrow score, i.e. the proportion of times that the stimulus was perceived with narrow strips in relief. The mean of eight observers (solid dots) was fitted with a raised sine wave (thick line) to extract the leftward bias for the assumed light source position (a rightward shift away from zero in the graph). The leftward bias can easily be seen by comparing the narrow scores for the  $-90$  and  $90^\circ$  stimulus orientations. The physiological results (bottom graph) show the N2 amplitude modulation with stimulus orientation. Error bars correspond to 95% confidence intervals [13].

sinewave whose phase (horizontal shift) allows a degree of freedom. Most observers presented a strong bias to the left for the illumination position, consistent with our previous findings [8]. The mean left light bias was  $13.8^\circ$  when the stimulus was presented centrally, a value significantly larger than zero ( $t(7) = 4.11$ ,  $p < 0.01$ ). This left bias was greatly reduced when the stimulus was presented para-foveally: when the stimulus was presented in the left visual field, the mean leftward light bias was small but significantly larger than zero ( $5.1^\circ$ ;  $t(7) = 2.12$ ,  $p < 0.05$ ) and at the same eccentricity but in the right visual field, the light bias was  $3.8^\circ$  (not significantly different from zero:  $t(7) = 1.03$ ,  $p > 0.1$ ).

**Physiological results:** We summarise below the main effects found from our ERP recordings. We present the analysis for amplitude measures using the following four time windows: P1 (96–104 ms), N1 (160–180 ms), P2 (232–248 ms), and N2 (280–300 ms). A four-way ANOVA of these data was applied using the factors eccentricity (stimulus position: left, centre, right), orientation (stimulus orientation: 16 levels), hemisphere (cerebral hemisphere: left, right), and site (lateral electrode site: 14 levels). Because we used an average reference, all reported effects are meaningful only in interaction with electrode site. However, provided that a significant overall effect was obtained, we also performed additional and more specific ANOVAs in which we grouped our electrodes into five areas: occipital (O1/2, PO9/10), parietal (P3/4, P7/8), temporal (TP9/10, P9/10), central (C3/4, T7/8), and frontal (F3/4, F7/8, FP1/2). Each of these groups includes an equal number of electrodes in the left and right hemispheres.

**Stimulus eccentricity:** Presenting our stimuli centrally or to the left or right of the fixation point produced dramatic changes in measured brain potentials. In particular, when all the lateral sites were considered together, P1 and N1 were significantly larger when the stimulus was presented foveally rather than para-foveally (eccentricity  $\times$  site,  $F(26,182) > 8.8$ ,  $p < 0.001$ ). This main effect of eccentricity was again present later for N2, ( $F(26,182) = 4.1$ ,  $p < 0.05$ ), but not significant for P2 ( $F < 0.6$ ). In addition, and as expected, the N1 and P2 amplitudes were larger on the contralateral side of the stimulus presentation (interaction eccentricity  $\times$  hemisphere  $\times$  site,  $F(26,182) > 3.7$ ,  $p < 0.05$ ), but the effect disappeared for N2.

When electrodes were grouped into five areas, we found a main effect of eccentricity in the P1 component only in the occipital, parietal, and frontal areas ( $F(2,14) > 5.9$ ,  $p < 0.05$ ) and in the N1 in all five areas ( $F(2,14) > 34.8$ ,  $p < 0.001$ ). No effect was found for the P2 and the N2 was significant only in the temporal area ( $F(2,14) = 8.7$ ,  $p < 0.05$ ).

A significant eccentricity  $\times$  hemisphere interaction was revealed in all five areas for the N1 component ( $F(2,14) > 4.6$ ,  $p < 0.05$ ), whereas this effect was significant for the P2 component only at parietal electrode sites ( $F(2,14) = 4.6$ ,  $p < 0.05$ ).

**Stimulus orientation:** The amplitude of the P2 and N2 was influenced by stimulus orientation (orientation  $\times$  site,

$F(195,1365) > 1.6, p < 0.05$ ). This can be seen by looking at individual ERP traces in Fig. 3.

Analysis of the grouped electrodes revealed that a main effect of orientation in both the P2 as well as the N2 was only present at occipital, temporal, and parietal areas ( $F(15,105) > 2.0, p < 0.05$ ). Looking more closely at the variation of the N2 amplitude with orientation, we found a modulation of the amplitude similar to the one observed behaviourally, with large biases for leftward illuminations (left bias equalled 27.2, 26.7 and 22.4 for occipital, temporal and parietal areas, respectively; see Fig. 2, bottom graph).

**Shape perception:** When the stimulus orientation produced a bi-stable percept, the response of the participant varied from trial to trial even though the very same stimulus was being presented. We now restrict our analysis to those orientations that produced the most ambiguous stimuli and define a new variable, orientation2, as the set of orientation pairs (90,105) and (270,285)°. For each pair orientation2, we split our recordings into trials where the observer responded seeing narrow strips in relief *vs* wide strips. Thus we performed a four-way ANOVA with the factors orientation2 (2 pairs), response (narrow, wide), hemisphere (left, right), and lateral electrode site (14 levels). We report here only the analysis for the P1 component and central stimulus presentation.

Analysis across all lateral electrode sites revealed a significant 3-way response  $\times$  hemisphere  $\times$  site interaction

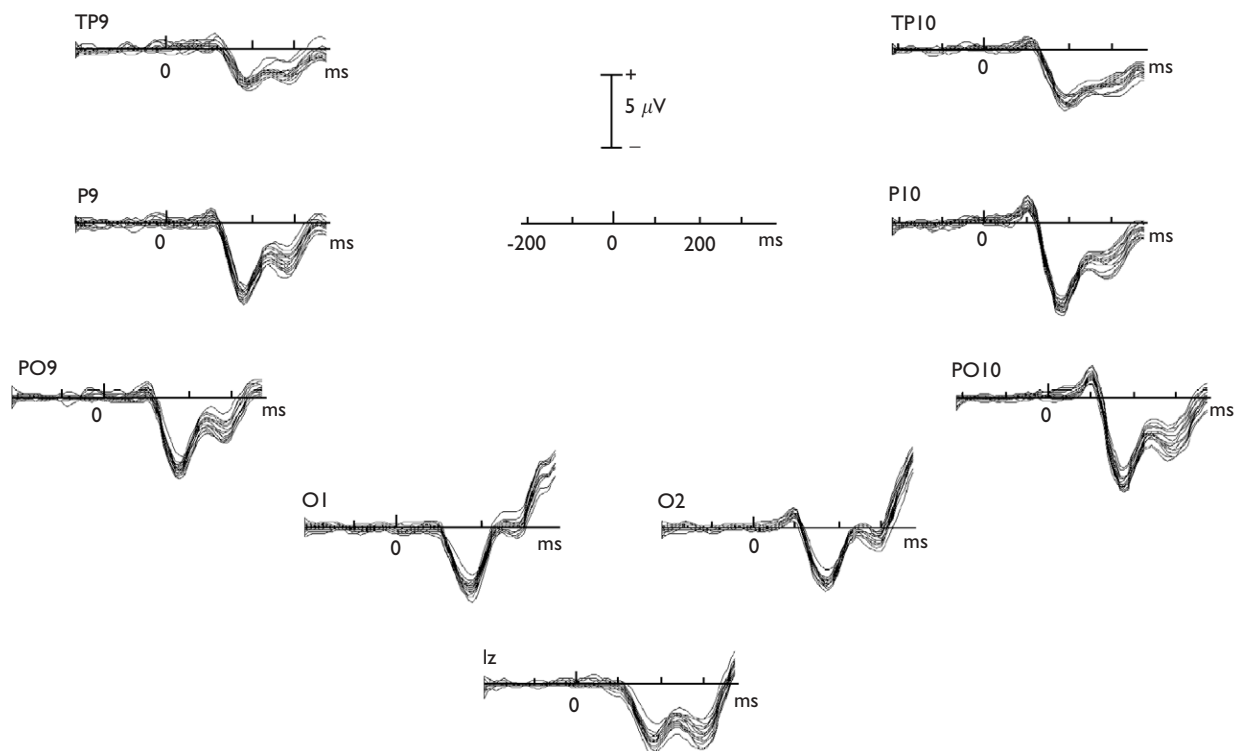
( $F(13,91) = 3.1, p < 0.05$ , which was further modulated by orientation2 ( $F(13,91) = 2.5, p < 0.01$ ).

Analysing different groups of electrodes separately, we found a significant hemisphere  $\times$  response interaction for the P1 time window in temporal and occipital areas ( $F(1,7) > 4.7, p < 0.05$ ). As illustrated in Fig. 4a, this effect was due to a larger hemispheric difference for the wide (temporal sites:  $M(\text{left vs right}) = -0.53 \text{ vs } 0.76 \mu\text{V}$ ; occipital sites:  $M(\text{left vs right}) = -0.06 \text{ vs } 1.20 \mu\text{V}$ ) than the narrow responses (temporal sites:  $M(\text{left vs right}) = 0.54 \text{ vs } 0.45 \mu\text{V}$ ; occipital sites:  $M(\text{left vs right}) = 0.32 \text{ vs } 0.68 \mu\text{V}$ ).

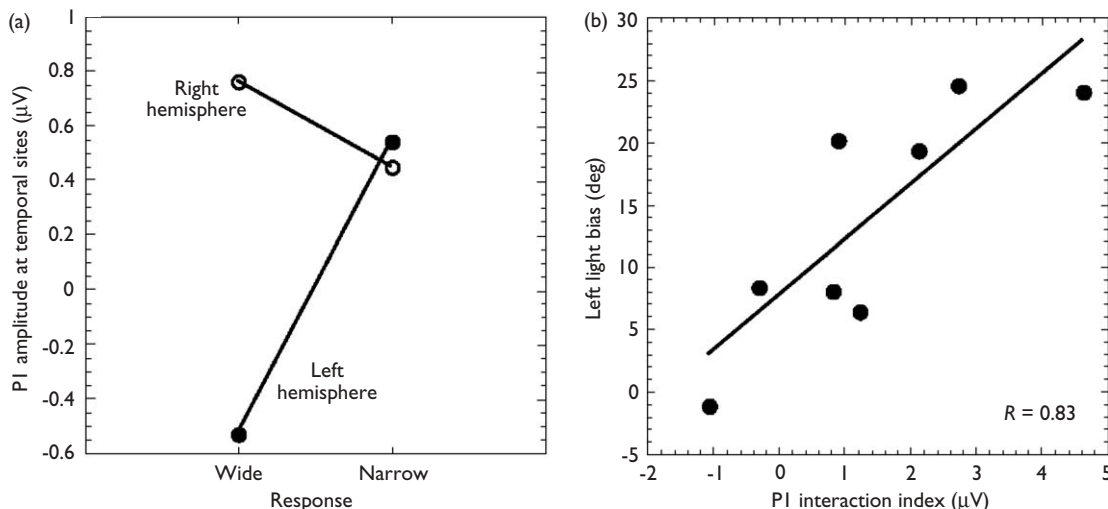
We were interested to test whether this interaction was related to the assumed light source position. To test this possibility, we defined the P1 interaction index as follows:  $\text{P1 interaction index} = [A(\text{left,wide}) - A(\text{left,narrow})] - [A(\text{right,wide}) - A(\text{right,narrow})]$ , where  $A(x,y)$  is the P1 amplitude at hemisphere  $x$  corresponding to  $y$  responses. Across all observers, we found a strong correlation of this P1 interaction index with the leftward bias for the light source position ( $R = 0.83$ ; Fig. 4b).

## DISCUSSION

The perception of shape from shading is ambiguous if the light source position is unknown. We have presented to human observers a shaded pattern whose perceived shape changes depending on its orientation. We found a robust bias to the above-left for the assumed light source position, very strong when the stimulus was presented foveally and



**Fig. 3.** Event-related potentials (ERPs). The ERPs are shown here for all 16 stimulus orientations across the occipital and temporal sites. Each trace within each site figure corresponds to a particular orientation. It can be seen that orientation had a clear effect on P2 and N2 components (see traces for P9/10 and PO9/10 sites), but little effect on P1 and N1 components.



**Fig. 4.** Relationship between PI and response for the most ambiguous stimulus orientations. (a) Interaction between perceived shape and the recorded hemisphere for the temporal area. (b) The interaction between perceived shape and hemisphere was correlated with the individual observers' leftward bias for the light source position.

weaker but still present when the stimulus was presented para-foveally.

While observers looked at the ambiguous shaded pattern, we recorded event-related potentials from their scalp. We found that early activation in human visual areas was correlated with the perception of shape from shading. More specifically, we found that the N2 component (about 300 ms after stimulus onset) in the occipital, temporal and parietal areas was modulated by stimulus orientation. This modulation was biased in a way that mimicked the leftward bias for the light source position found behaviourally. In addition, we found that the P1 component (about 100 ms after stimulus onset) in occipital and temporal areas was correlated with perceived shape and individual observers' bias for the light source position.

The interaction of the P1 amplitude with the participants' response indicates that the shape of the shaded pattern has been disambiguated within the first 100 ms of stimulus presentation. Moreover, the P1 component is believed to emerge from Brodmann's areas 18 and 19 [9]. Recent imaging studies also showed activity in area 17 of the occipital lobe [4,5]. Taken together, these results indicate that shape-from-shading is first computed in the occipital areas with little contribution from higher cortical areas. The interpreted three-dimensional structure is likely to be analyzed further by temporal areas [10,11] and interact with other interpretations from other depth cues in parietal areas [1].

If shape-from-shading is mostly computed in the occipital areas, then the assumed light source position should also be represented early on in the visual system. Because occipital

areas are retinotopic, any representation at that level should also be retinotopic. A direct prediction is that the bias for the light source position should be retino-centric rather than world-centric. Psychophysical evidence supports this prediction [12].

In summary, we found an early neural response that was correlated with the perception of ambiguous shaded objects. These results indicate that an assumed light source position is represented early on in the visual system to enable a quick first appreciation of the scene. This interpretation suggests a bottom-up mechanism for shape from shading.

## REFERENCES

1. Taira M, Nose I, Inoue K and Tsutsui KI. *Neuroimage* **14**, 959–966 (2001).
2. Tsutsui KI, Jiang M, Yara K *et al.* *J Neurophysiol* **86**, 2856–2867 (2001).
3. Tsutsui KI, Sakata H, Naganuma T and Taira M. *Science* **298**, 409–412 (2002).
4. Humphrey GK, Goodale MA, Bowen CV *et al.* *Curr Biol* **7**, 144–147 (1997).
5. Lee TS, Yang CF, Romero RD and Mumford D. *Nature Neurosci* **5**, 589–597 (2002).
6. Ramachandran VS. *Nature* **331**, 163–166 (1988).
7. Sun J and Perona P. *Nature Neurosci* **1**, 183–184 (1998).
8. Mamassian P and Goutcher R. *Cognition* **81**, B1–B9 (2001).
9. Di Russo F, Martinez A, Sereno MI *et al.* *Hum Brain Mapp* **15**, 95–111 (2001).
10. Connor CE. *Curr Biol* **11**, R776–R778 (2001).
11. Hanazawa A and Komatsu H. *J Neurosci* **21**, 4490–4497 (2001).
12. Howard IP, Bergström SS and Ohmi M. *Perception* **19**, 523–530 (1990).
13. Loftus GR and Masson MEJ. *Psychon Bull Rev* **1**, 476–490 (1994).

Acknowledgements: This work was supported in part by Human Frontier Science Program grant RG0109/1999-B.