seen in polar projection (the configuration that normally elicits the most rigid percept) would be seen as deforming.

Hence Sinha and Poggio's results clearly demonstrate that future psychophysical investigation on the perception of 3D shapes will have to take into account learning processes that can take place on relatively short time scales. More generally they also lead to reconsider classical schemes of 3D shape perception in terms of: (1) the type of object representation involved in visual processes, and (2) the existence of top-down control of these processes.

References

1 Wallach, H. and O'Connell, D.N. (1953) The kinetic depth effect *J. Exp. Psychol.* 45, 205–217

2 Sinha, P. and Poggio, T. (1996) Role of learning in three-dimensional form perception *Nature* 384, 460–463

3 Ittelson, W.H. (1960) Size, shape, perspective, in *Visual Space Perception*, p. 80, Springer

4 Julesz, B. (1971) Cyclopean perception in perspective, in *The Foundation of Cyclopean Perception*, pp. 300–314, The University of Chicago Press

5 Longuet-Higgins, H.C. and Prazdny, K. (1980) The interpretation of a moving retinal image *Proc. R. Soc. London Ser. B* 208, 385–397

6 Ullmann, S. (1979) The interpretation of structure from motion, in *The Interpretation of Visual Motion*, pp. 133–175, MIT Press

7 Poggio, T., Torre, V. and Koch, K. (1985) Computational vision and regularization theory *Nature* 317, 314–319

8 Ames, A. (1951) Visual perception and the rotating trapezoidal window *Psychol. Monogr.* 65, 1–32

9 Mingolla, E. and Todd, J. (1981) The rotating square illusion *Percept. Psychophys.* 29, 487–492

10 Cornilleau-Pérès, V., Marin, E. and Droulez, J. (1996) The dominance of static depth cues over motion parallax in the perception of surface orientation *Perception* 25 (suppl.), 40

Response from Sinha and Poggio

P. Sinha and T. Poggio are at E25-201, Center for Biological and Computational Learning, Department of Brain and Cognitive Sciences, MIT, 45 Carelton Street, Cambridge, MA 02142, USA.

tel: +1 617 253 0547
fax: +1 617 253 2694
e-mail: sinha@ai.mit.edu

Lamouret, Cornilleau-Pérès and Droulez raise a number of very interesting points in their **Comment** article. We should like to take this opportunity to emphasize one of these issues that we find particularly important and intriguing but could not dwell upon adequately in the original paper, for lack of space.

Lamouret *et al.* remark on how learning initially depends on bottom-up sensory-information processing that uses generic biases such as those favoring object rigidity. However, the learning subsequently can overwhelm the results of such bottom-up processing. The percept, apparently, is controlled to different extents at different times by the generically processed sensory information on the one hand and object-specific learned expectations on the other. The big question is: How does the brain strike a compromise between sensation and, for want of a better term, hallucination? The parameters determining the relative contributions of the two quantities to the overall percept are likely to be a function of time in two ways. (1) Expectations will exercise greater control in determining percepts the longer the training time. (2) The bottom-up sensory information will become increasingly evident the greater the stimulus inspection time. The well-known hollow-mask illusion serves as a nice illustration of this point. The illusion often persists even under binocular viewing. If one subscribes to the accounts of the illusion that are based on familiarity, then it is reasonable to suggest that the greater the familiarity of an observer with faces, the more susceptible the observer will be to perceiving the illusion. On the other hand, the longer one binocularly inspects the hollow mask, the more likely one is to perceive its correct (hollow) structure. Our experimental results follow a similar pattern. The key question that needs to be addressed to explain these empirical observations is how expectations are combined with sensory information to yield the overall percept. It seems to be a rather involved question, given that the combination strategy is a function of at least two temporal variables. Among others, it is likely to prove interesting to colleagues who have been studying so fruitfully the issue of cue-combination, except that one of the cues would now be 'internal' to the visual system. Work on this problem holds the exciting potential of bringing together two big, and so far largely independent, streams of research – one examining 'bottom-up' processing and the other 'top-down' strategies.

On another issue, Lamouret *et al.* correctly point out that the shape representation schemes we discuss are better characterized as implicit versus explicit, with emphasis on the nature of the coded variables. The visual system might also possess some limited ability to extract viewer-centered depth information, which, though an 'explicit' encoding of shape, cannot readily be subjected to arbitrary projectional transformations.

Lamouret *et al.* deserve thanks for summarizing our results so clearly and for suggesting and highlighting some of the important issues that need to be tackled next.

Higher-order processes in auditory-change detection

Risto Näätänen and Kimmo Alho

R. Näätänen and K. Alho are at the Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, PO Box 13 (Meritullinkatu 1), FIN 00014, Finland.

tel: +358 9 1912 3445
fax: +358 9 1912 2924
e-mail: risto.naatanen@helsinki.fi

The paper by Schröger and Wolff[1] is, perhaps, the first study that has clearly succeeded in demonstrating what is memory-related and what is memory-unrelated (as we interpret the results) in the enhancement of an electric brain response to an infrequent stimulus change. In this study, a sound (the 'standard') with a certain apparent location (manipulated by the interaural time difference) was repeated at short intervals, and was occasionally replaced by an identical sound, which had a slightly different apparent location (the 'deviant'), whilst the attention of the subject under investigation was directed elsewhere. These deviants elicited an event-related potential (ERP), which was enhanced relative to that elicited by the standard. This enhancement emerged as a negative shift, at the time region of 100–250 ms from stimulus onset, in the deviant–standard difference wave.

To account for this enhancement, firstly one needs to consider the fact that the sound-location specific afferent

 PII: S1364-6613(97)01013-9