

Visual Perception: Shaping What We See Dispatch

David A. Leopold

A new psychophysical study has examined the free flow of perception as observers viewed stimuli with several possible visual interpretations. The results suggest that our subjective impression of such ambiguous patterns may be more closely linked to the brain’s encoding of complex shape than previously appreciated.

It is interesting to consider that the first thinkers who pondered visual perception were inclined to imagine the eye as an active organ, shooting out streams of light at the surroundings to ‘feel’ the structure of objects and scenes. This *emanation* theory of sight was not a fringe belief held only by mystics, but was the accepted view among such great minds as Plato, Euclid and Ptolemy [1]. In contrast, modern theories of vision begin with an image of the world, cast innocently and passively onto the back of the eye. This image ceases to be an impartial physical entity only through its reception by the retina and brain, which perform a rigorous analysis aimed at understanding the structure and arrangement of objects from which the light pattern originated.

In the past few decades, microelectrode recordings in monkeys have provided great insight into this analytic aspect of vision [2]. The secret appears to lie in a highly organized neural network of explicit stimulus representations, acting first to dissect a visual scene into its most basic attributes, and later reconstituting it as understandable objects and forms. A prominent part in the initial dissection is played by the primary visual cortex, a large area replete with neurons responding to specific visual ‘primitives’ in the image, such as spatially localized edges or color patches. Higher-order representations, thought to derive from combinations of such primitives, are found at later cortical processing stages such as the inferotemporal cortex, where individual cells respond to geometrical shapes, objects, and even faces [3–6].

The visual cortex, as a whole, thus imposes an elaborate and massively parallel analysis upon the light image striking the retina, simultaneously encoding a diversity of low-level and high-level image features. It may be slightly alarming to consider that much of this sensory machinery can perform its task, at least to a first approximation, even when an animal is under general anesthesia. But, while the ultimate link with perception remains poorly understood, this analytic aspect of vision is generally taken to be the core of the brain’s approach to the problem of vision,

allowing us to decode, understand, and ultimately interact with the visual surroundings.

But is it really possible that our visual perception of the world derives from an automatic and largely passive re-representation of the retinal image? Clearly, the ancients understood that vision consists of ‘looking’ in addition to ‘seeing’ — we do not simply sense our surroundings, but are actively engaged in them. Our brain is able to orient to particular elements of a scene, ignore others, and generate hypotheses about the structure of objects based on imperfect or incomplete data [7]. For some patterns that are inherently ambiguous, such active elements of perception are revealed through the brain’s continual exploration of alternative visual solutions. Typically, such patterns are ‘bistable’, offering exactly two percepts that alternate every few seconds in their dominance. They are often used to illustrate the interpretive nature of perception, as they provide the brain with freedom to ‘organize’ its input in accordance with structural properties of the stimulus [8], as well cognitive factors such as attention [9] or prior experience [10]. In modern-day speak, mechanisms that aide or guide perceptual organization with these or any stimuli are labeled ‘top-down’ effects — a term so often used for such a wide range of phenomena that it has virtually no explanatory power. A great challenge for vision scientists is to gain a more refined notion of how basic feature representations in the visual cortex interact with loftier parts of the brain involved in attention, abstraction, memory and decision-making.

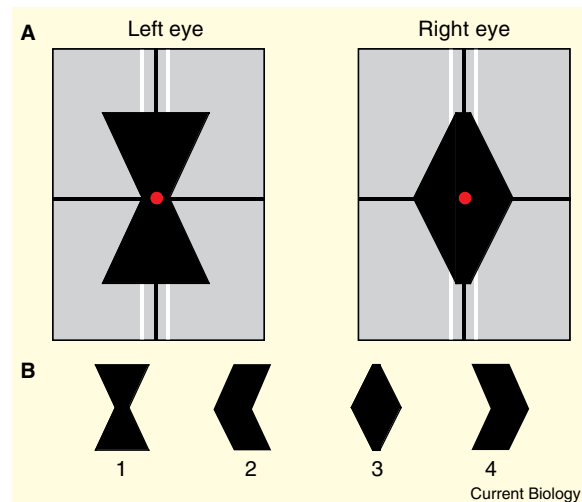


Figure 1. Example of stimulus used in the perceptual trapping experiment.

If the two patterns in (A) are presented stereoscopically to the two eyes, perception fluctuates between four possible coherent percepts (B). However the probability of seeing a particular stimulus at each point in time is not random, but instead highly dependent on the previous perceptual state.

A recent study by Suzuki and Grabowecky [11] may shed some new light on this difficult question. They explored how visual perception, when faced with an unusually ambiguous problem, ultimately settles upon an answer. Specifically, they developed a novel family of multistable patterns that, when viewed stereoscopically, each gave rise to four distinct shapes (Figure 1). A key to this study was the multiplicity of solutions, which allowed the authors to explore not only the timing of perceptual reversals, but also ‘where’ perception went at each point in time. Interestingly, they found that it did not jump randomly between all solutions, which might be expected given that each solution was seen, on average, for roughly equal amounts of time. Instead, perception tended to become ‘trapped’ between particular pairs of shapes.

With the stimulus in Figure 1, for instance, perception was much more likely to fluctuate between shapes 1 and 3, or 2 and 4, than any of the other combinations (Figure 2). This favoring of particular transitions, or *path dependence* as Suzuki and Grabowecky [11] call it, adds a new dimension to the investigation of multistable perception, and has potentially important implications for natural vision as well. Why might the brain, faced with several equally valid solutions to a visual problem, pursue certain paths through an abstract ‘shape-space’ at the expense of others?

To get at this point, Suzuki and Grabowecky [11] drew upon previous results that explored the neural encoding of global shape using a novel aftereffect [12,13]. In those studies, Suzuki and colleagues used the psychophysical technique of adaptation to provide evidence that basic shapes may be represented in a reciprocal or ‘opponent’ manner in the visual system. Opponent encoding schemes are common in vision. They are known to underlie the analysis of first-order stimulus attributes such as luminance and color [14], and recent evidence suggests that they may play a role in the representation of much higher stimulus attributes as well [15]. For geometrical patterns, opponent principles could aid greatly in the analysis of global shape attributes such as convexity or asymmetry, as they offer a number of computational advantages over classic ‘feature detectors’ that simply report a presence or absence of a particular shape. Interestingly, Suzuki and Grabowecky [11] discovered that the trapping of perception occurred primarily between shapes that had been previously identified as opponent pairs. This subtle point is critical to understanding the implications of their results, as it provides a direct link between the observers’ subjective experience and the neural representations in the visual cortex.

Perceptual trapping demonstrates that even when the brain is given great freedom to interpret its sensory inputs, it does not wander randomly through solutions, nor does it seek them based upon abstract or semantic properties of stimuli. Instead, it appears to be bound by principles stemming directly from the underlying neural encoding scheme — in this case the opponent coding of global shape. As usual, the ultimate interpretation of these observations depends upon existing preconceptions. Opinions regarding the nature

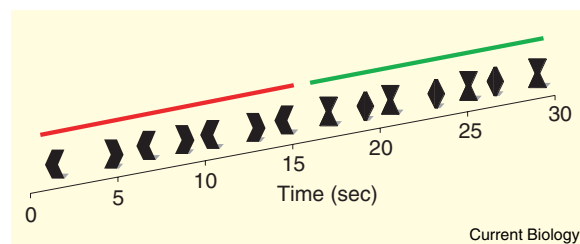


Figure 2. Example of perceptual trapping with the stimulus shown in Figure 1.

The percept of a subject is shown to alternate as a function of time (despite the unchanging visual stimulus). While each of the four visual solutions commonly reaches perceptual dominance, the majority of subjective transitions are between opponent shape pairs, and perception is thus ‘trapped’ (red and green highlighted regions).

of perceptual multistability remain diverse [16–20], and those who take it to be fundamentally a manifestation of instability in the sensory domain will likely embrace these findings as further evidence for that viewpoint. Yet, it may be of great value to consider these results not only in terms of local coding, but also with regard to interactions between brain areas.

Much evidence suggests that the resolution of sensory ambiguity is aided or even steered by high-level, executive centers in a kind of perceptual ‘exploration’ [18]. From this vantage point, the present findings might give new insight into the level of visual processing at which such external signals intervene. One of many possibilities is that higher centers communicate directly and specifically with shape-selective neurons in areas like the inferotemporal cortex, prompting them to issue a general ‘reevaluation’ of the sensory input that, in turn, might affect many levels of visual representation.

Regardless of which areas may be responsible for initiating and/or interpreting the perceptual changes, these findings make the important new point that the course of subjective perception, even when given the possibility to wander freely between equally valid interpretations of a stimulus, is bound to the basic principles of sensory encoding in the visual cortex.

References

- Howard, I.P. and Rogers, B.J. (2002). Seeing in depth. Ontario, Canada, I. Porteus.
- Hubel, D. (1988). Eye, Brain and Vision. (Scientific American Library).
- Desimone, R., Albright, T.D., Gross, C.G. and Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* 4, 2051–2062.
- Janssen, P., Vogels, R. and Orban, G. A. (2000). Selectivity for 3D shape that reveals distinct areas within macaque inferior temporal cortex. *Science* 288, 2054–2056.
- Logothetis, N.K. and Sheinberg, D.L. (1996). Visual object recognition. *Annu. Rev. Neurosci.* 19, 577–621.
- Perrett, D.I., Rolls, E.T. and Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Exp. Brain Res.* 47, 329–342.
- Gregory, R. (1997). Eye and brain: The psychology of seeing. (5th edn), (Princeton University Press).
- Koffka, K. (1935). Principles of Gestalt Psychology. New York, Harcourt, Brace and World.
- Suzuki, S. and Peterson, M.A. (2000). Multiplicative effects of intention on the perception of bistable apparent motion. *Psychol. Sci.* 11, 202–209.

10. Girgus, J.J., Rock, I. and Egatz, R. (1977). The effect of knowledge of reversibility on the reversibility of ambiguous figures. *Percept. Psychophys.* 22
11. Suzuki, S. and Grabowecky, M. (2002). Evidence for perceptual 'trapping' and adaptation in multistable binocular rivalry. *Neuron* 36, 143–157.
12. Suzuki, S. (2001). Attention-dependent brief adaptation to contour orientation: a high-level aftereffect for convexity? *Vision Res.* 41, 3883–3902.
13. Suzuki, S. and Cavanagh, P. (1998). A shape-contrast effect for briefly presented stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1315–1341.
14. Gegenfurtner, K.R. and Scharpe, L.T. (1999). *Color vision: from genes to perception.* (Cambridge University Press).
15. Leopold, D.A., O'Toole, A.J., Vetter, T. and Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nat. Neurosci.* 4, 89–94.
16. Rock, I. (1984). *Perception.* (Scientific American Library).
17. Long, G.M., Toppino, T.C. and Mondin, G.W. (1992). Prime time: fatigue and set effects in the perception of reversible figures. *Percept. Psychophys.* 52, 609–616.
18. Leopold, D.A. and Logothetis, N.K. (1999). Multistable phenomena: changing views in perception. *Trends Cognit. Sci.* 3, 254–264.
19. Blake, R. and Logothetis, N.K. (2002). Visual competition. *Nat. Rev. Neurosci.* 3, 13–21.
20. Tong, F. and Engel, S.A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature* 411, 195–199.