



ELSEVIER

# Perspectives on science and art

## Bevil R Conway<sup>1</sup> and Margaret S Livingstone<sup>2</sup>

Artists try to understand how we see, sometimes explicitly exploring rules of perspective or color, visual illusions, or iconography, and conversely, scientists who study vision sometimes address the perceptual questions and discoveries raised by the works of art, as we do here.

### Addresses

<sup>1</sup>Neuroscience Program, Wellesley College, Wellesley, MA 02481, USA

<sup>2</sup>Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA

Corresponding author: Conway, Bevil R ([bconway@wellesley.edu](mailto:bconway@wellesley.edu))

**Current Opinion in Neurobiology** 2007, **17**:476–482

This review comes from a themed issue on  
Sensory systems  
Edited by Peter Mombaerts and Tony Zador

Available online 11th September 2007

0959-4388/\$ – see front matter  
Published by Elsevier Ltd.

DOI [10.1016/j.conb.2007.07.010](https://doi.org/10.1016/j.conb.2007.07.010)

The integration of visual art with the experimental study of vision has its roots in formal analysis of painting, but this approach has gone out of vogue in art-historical circles because it ignores the impact of cultural context on art appreciation. Advances in our understanding of how the brain works have resuscitated interest in linking art with vision science. Here we review some of these new links, with the caveat that we will not address visual aesthetics, but rather some nuts and bolts of making and looking at art, in an endeavor to enrich our understanding of art ‘in much the same way as a knowledge of bones and muscles has for centuries enhanced the ability of artists to portray the human body’ [1]. Some recent studies may even illuminate why cultural context is so important in visual art.

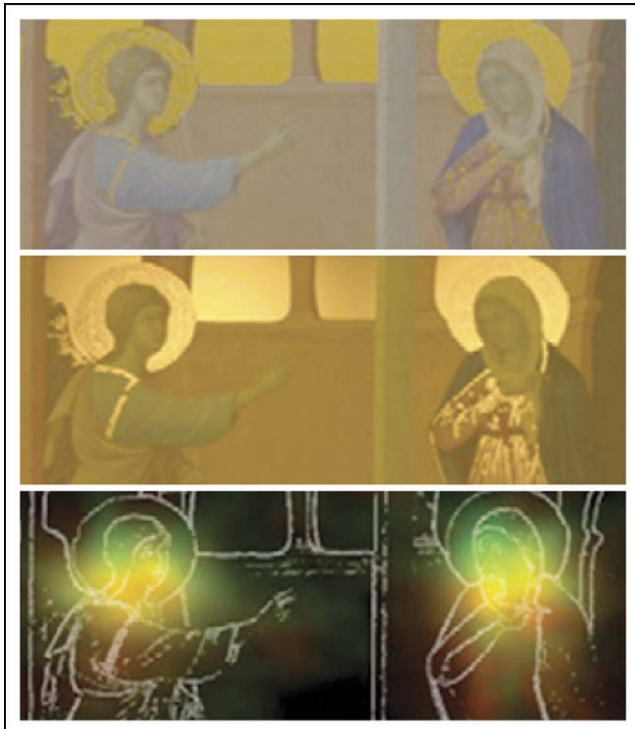
When we look at the world, or at a work of art, our eyes, and our visual attention, are constantly and alternately moving and fixating [2]. Our fixations are not randomly distributed across a scene, but rather concentrate on key regions—to a first approximation, on the areas of high local contrast [3]. The visual machinery that directs the eyes where to look can have two impacts on art. On the one hand, artists have developed techniques to direct your gaze; on the other hand, the unconscious machinery directing the gaze of the artist may influence which parts of the scene the artist portrays or emphasizes. Leonards *et al.* [4] have argued that Renaissance artists used gold for its special reflective quality under candlelight to control

the viewers’ gaze. The authors argue that ‘the glow of the gold induced shifts in fixations to symbolically important regions of the painting’ (Figure 1). Candle illumination introduces areas of local contrast that is not apparent under daylight conditions—the thin lines of gold extending into the bodies of the two subjects. It is these regions that show increased fixations under low illumination, raising the possibility that the viewer’s gaze and attention are drawn to these regions not because of their symbolism, as Leonards *et al.* argue, but rather because of their low-level visual salience.

Eyes are especially potent in attracting the gaze, both because of their behavioral significance and because they are the regions of high local contrast. Tyler [5•] examined 265 portraits spanning the past 600 years and found a strong tendency for one eye to be centered on the vertical midline of the painting (Figure 2). A recent critique of Tyler’s observations asserts that ‘one eye tends to be relatively close to the vertical midline because of geometric constraints on the placing of a relatively large object, the head, within a pictorial frame’ suggesting that there is no behavioral significance to eye-centering [6•]. But a careful analysis shows that heads in portraits do not tend to be centered within the picture frame, but rather slightly offset [5•,7] (Figure 2, bottom panel). The bimodal distribution of nose position indicates that centering one eye has a stronger influence than predicted from simply centering the head (Figure 2, bottom panel). Eye-centering may have its roots in ‘hidden principles [that] are operating in our aesthetic judgments,’ as argued by Tyler, or it may simply reflect the fact that the artists themselves are fixating on one eye (because you can only look at one eye at a time) and placing it in the center of the canvas because, absent any other visual cues (as the blank canvas appears to the artist), the center is where we look. Moreover, the centered eye is usually the nearer of the two and therefore the most salient.

The tendency for eye-centering has become less pervasive in the 20th and 21st centuries [6•]. A random selection of Picasso’s portraits, for example, contains only one with a centered eye (the bottom one in Figure 3). Following the artistic imperative of continuous innovation, avant-garde movements, like those led by Picasso, sought for ever new modes of expression and generated a self-conscious break not only with traditional modes of representation but also with traditional methods of production. Artists have become increasingly aware of (and in some cases averse to) the human constraints on art making; as a result, some contemporary artists have attempted to delineate forms according to pre-determined formulas, as in Sol

Figure 1

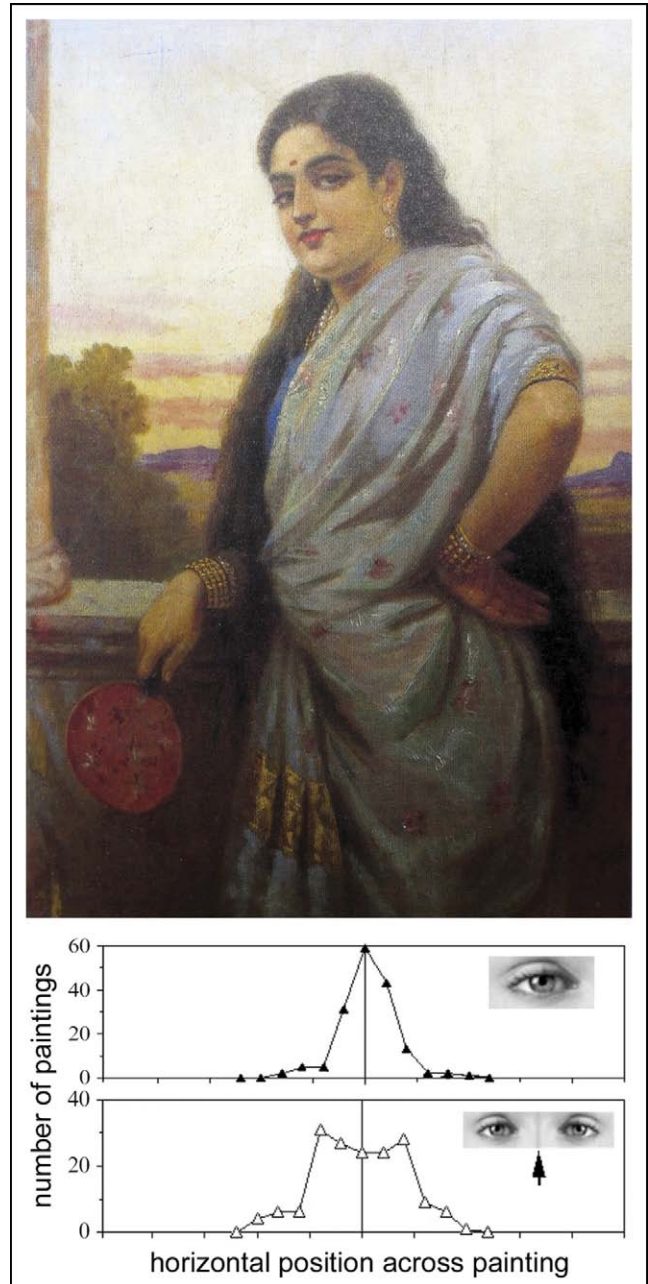


*The Annunciation* (1311) by Duccio in daylight (top) and candlelight (middle). Bottom shows fixation regions in daylight (green) and candlelight (red). From reference [4].

Lewitt’s instructions for wall drawings, Roxy Paine’s sculpture machines, which extrude impersonal factory-made objects, and Yoko Ono’s Sky TV, which is a movie generated by training a camera on a pre-determined plot of sky. This kind of work is often invested with a heavy conceptual weight, though perhaps not surprisingly, many find it aesthetically wanting and question its claim to art [8].

Because we have high acuity for only the small region of the visual field occupied by the fovea, in order to generate the perception that the entire world in front of us is seen at high resolution, we constantly shift our eyes from one location to the next, maintaining in visual memory a series of retinal snapshots. The brain’s representation of eye proprioceptors was recently discovered [9], raising the possibility that humans could use proprioceptive information from the accommodation muscles of the eyes as a rough gauge of depth, like chameleons do [10]. Such proprioceptive information may be part of the ‘muscle memory’ that artists describe as so important in image making. Of course, there are other cues to depth, one of which, motion parallax, requires movement of the viewer’s head and another, stereopsis, requires precise alignment of the two eyes. Neither of these depth cues is available to the artist who sets out to make a flat static image. Artists must exploit other monocular cues, like

Figure 2



*Portrait of a Woman* (1895) by Raja Ravi Varma. Bottom panels quantify the eye-centering (top) and bi-modal head distribution (bottom) in 600 years of portraits. Bottom panels from references [5\*,7].

perspective, shading, and occlusion. People who lack stereopsis may actually have better access to these monocular depth cues: A survey we conducted shows that artists, including Rembrandt, are more likely to be strabismic than non-artists [11,12].

One consequence of having the eyes, and even the head, in constant motion is that one-point perspective is not



Figure 3

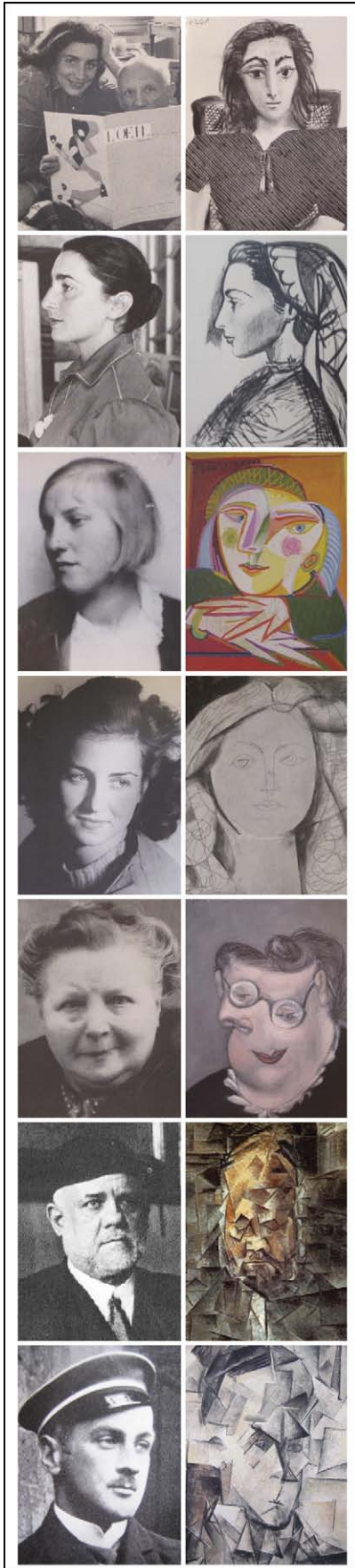
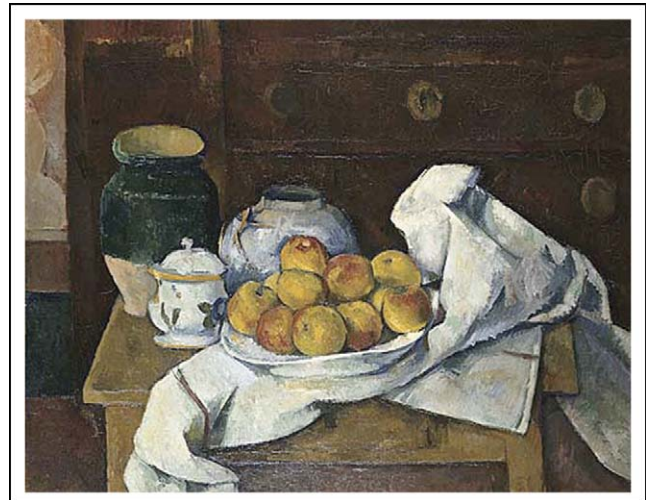


Figure 4



*Still Life with Commode* (1887–1888) by Paul Cezanne.

maintained. This may account for the technique of certain artists to generate images that lack a coherent single-point perspective. Cezanne's still lifes often depict different regions and often different parts of the same object, from different perspectives. The green jar in his *Still Life with Commode* (Figure 4) is an example: In reality, one would have to be looking down on the scene in order to see the opening of the jug as an ellipse; from this perspective, the bottom of the jug should be more circular than the top, not less as Cezanne depicts it. As with the eye-centering, one wonders whether this is a conscious or unconscious aesthetic trope, or whether it simply comes as a byproduct of painting from life, where the artist's eyes and head are in constant motion, governed by regions of interest and biased by meaning, unconstrained by the 'correctness' in painting advocated by Leonardo da Vinci [13]—the jug sits securely on the table, so the bottom of it must be flat; we pour liquids out of the jug, so the top must be round.

A recent study found that autistic children are not subject to the same cognitive biases: they depict a circle viewed from the side accurately, as an ellipse, while age-matched controls tend to depict it as more circular than it should be [14] (see also references [15,16,17]). That is, normal children impose on their drawings their knowledge about the three-dimensional shape of the object, rather than reproducing the true flat projection the object casts on their retinas, whereas autistics can sometimes draw

**(Figure 3 Legend)** Several portraits by Picasso (right), and photographs of each subject (left): From top: Jacqueline Picasso (1957), Jacqueline Picasso (1955), Marie-Thérèse Walter (1936), Françoise Gilot (1946), Emilie Marguerite Walter (1939), Ambroise Vollard (1910), and Wilhelm Uhde (1910).

three-dimensional objects remarkably accurately. Autistics' drawings are typically of non-biological subjects, and usually mechanical, rich in perspective. Nevertheless, these drawings rarely abide by a single-point perspective, showing instead multiple perspectives, each consistent for only a local region or object in the scene [18<sup>\*</sup>], a scale that may match, and be governed by, the size of the region of the scene that is in the center of gaze at one time. What we find astonishing about these drawings is that they, like Cezanne's painting, can look quite accomplished—we are not put off by the lack of single-point perspective. Some contemporary artists, aware that the constraint of single-point perspective is unnecessary for image recognition, make images that flagrantly violate the expected one-point perspective rules. David Hockney, for example, has pushed the idea of multiple perspectives in a single image to an extreme, depicting a chair that is still recognizable and exists in depth (Figure 5).

That Cezanne's jug and Hockney's chair are still recognizable support Cavanagh's arguments that the rules that govern vision do not adhere to those of traditional physics, but rather an 'alternative' physics, which can reveal fundamental features about how our visual systems func-

tion [19<sup>\*\*</sup>]. Our tolerance of multiple points of view in a single painting suggests that one characteristic of this alternative (psycho) physics is that depth cues are processed locally; that is, it is not necessary for the perspective cues to be consistent across the entire image. This idea is consistent with the plausibility at first glance of impossible figures, like Penrose's triangle, and with observations from psychophysics and neurophysiology that figure-ground segregation, depth ordering, and surface stratification are processed at early stages in the visual system, where receptive fields are small and processing is necessarily local [20–27].

Artistic depictions of shadows, transparency, and reflections are other examples of Cavanagh's alternative physics (Figure 6). Depth-from-shading cues need not be consistent across an entire image. Shadows are conspicuously absent from art until relatively recently [28], but even when artists began to depict them, the shadows generally would not arise under physically plausible lighting conditions. Therefore, our brains do not rely on the physics of the real world to interpret shadows [29]. Furthermore, our brains do not even require the wavelength composition of shadows to be consistent with plausible illumination conditions: We simply require local regions to be relatively darker, and of relatively desaturated color [30]; in fact, the shadow itself can be of any color as long as these rules are met, as the Fauvists showed, and Kingdom *et al.* recently confirmed [31].

One reason why shadows, perspective, and reflection may not need to be physically accurate for an acceptable impression of depth, transparency, or shine is that the visual system evolved under conditions in which these features are not stable and not consistent within a scene over time: Shadows change as daylight changes (hence the inconsistencies of shadows in impressionist paintings, which were often painted outdoors over a stretch of time during which the sun and shadows move); similarly, perspective and reflections change second to second as we move our eyes across a scene. Therefore, there would have been little biological benefit to incorporate the rules for global perspective or illumination into our visual computations. Furthermore, since the world is generally lawful, it would be unnecessary to have mechanisms for rejecting possibilities that are inconsistent with the laws of physics. Instead, our interpretation of the three-dimensional organization of a scene is generated by stitching together multiple impressions, each localized to the region centered on our gaze direction. The fragmented nature of seeing makes the discovery of single-point perspective, and the subsequent obligation of paintings to conform to it, a remarkable demonstration of the power of connoisseurship and of the importance of cognitive function (learning) in aesthetic appreciation. Rules like one-point perspective must be learned, and they do not even require visual experience, as shown by the ability of a

Figure 5



Chair Jardin de Luxembourg (1985) by David Hockney.



Figure 6



(Left) Colliding shadows in Lorenzo di Credi's *Annunciation* (1480–1485). (Right) Impossible reflection in Vermeer's *Young woman with a water jug* (ca. 1662).

congenitally blind artist to generate perspective drawings complete with accurate foreshortening [32]. The importance of learning is, of course, not restricted to visual aesthetics, as the example of tonal unity in music shows: Naïve listeners are hard-pressed to notice (or care) if a piece of music begins and ends in a different key, but since its discovery, coincidentally around the same time as perspective, tonal unity has become an important ordering principle in music [33]. By analogy with perspective, which is effective on small local spatial scales, the experience of tonal closure is 'restricted to fairly short time spans' [34].

As we move our eyes over an image, not only is our gaze direction non-random, but consequently so is our attention and our information intake. Therefore, those parts of an image that are more salient generate more neuronal activity, not only in low-level visual areas but also in higher level regions of our visual system, regions concerned with object representation and memory. Some parts of an image are more strongly represented in our brains: object versus background, face versus torso, figure versus tree. Drawings often reflect these mental representations as biases, or even inventions. For example, naïve subjects often draw objects with no background, 3-D objects to show more than would be visible from a single viewpoint, faces as round with no forehead, smiles as curving upward, and children with the proportions of miniature adults. These representations probably reflect efficiency in the way in which our brains represent just behaviorally relevant aspects of our visual worlds. The fact that such mental representations can affect the structure of a drawing, even one made using direct observation, shows the importance of feedback

connections from high-level cognitive areas, back to lower level stages of visual processing, shaping our perceptual experience. Recent physiological studies have begun to explore the ways in which top-down rules generated through experience can alter the response properties of neurons at various stages of visual processing [35,36] and sculpt gross patterns of brain activity [37–39]. What restricts the set of features or categories to which we can become expert is yet unresolved. Comparative studies of art produced by different cultures may shed some light on the universality of iconography—lines, for example, are used in every culture to represent contours. Such studies may in turn shed light on the nature of neural plasticity.

Certain types of expertise, like our ability to recognize faces, are astonishingly good considering the difficulty of the task; humans and macaques have developed specialized neural machinery for detecting and recognizing faces. Such target selectivity must be generated via hierarchical stages of neural processing, beginning in the retina with a sensitivity to discontinuities, and then in V1 with a sensitivity to contour orientation and so on [40]. Recent physiological discoveries at an intermediate stage of visual form processing, in area V4, before the neural representation of faces, have shown how these basic elements are combined, establishing a sensitivity to curvature and to feature combinations [41]. In the same way we can understand the significance of lines in art in terms of the orientation selectivity of V1 neurons, vision scientists will probably uncover guiding principles in the establishment of more complex forms in V4 and how these can account for artistic representations of shape. A recent study has shown that late Bronze Age wall paintings are constructed from a limited



9. Wang X, Zhang M, Cohen IS, Goldberg ME: **The proprioceptive representation of eye position in monkey primary somatosensory cortex.** *Nat Neurosci* 2007, **10**:640-646.
10. Harkness L: **Chameleons use accommodation cues to judge distance.** *Nature* 1977, **267**:346-349.
11. Livingstone MS, Conway BR: **Was Rembrandt stereoblind?** *N Engl J Med* 2004, **351**:1264-1265.
12. Marmor MF, Shaikh S, Livingstone MS, Conway BR: **Was Rembrandt stereoblind?** *N Engl J Med* 2005, **352**:631-632.
13. Ono H, Wade NJ, Lillakas L: **The pursuit of Leonardo's constraint.** *Perception* 2002, **31**:83-102.
14. Ropar D, Mitchell P: **Shape constancy in autism: the role of prior knowledge and perspective cues.** *J Child Psychol Psychiatry* 2002, **43**:647-653.
15. Iarocci G, Burack JA, Shore DI, Mottron L, Enns JT: **Global-local visual processing in high functioning children with autism: structural vs. implicit task biases.** *J Autism Dev Disord* 2006, **36**:117-129.
16. Sheppard E, Ropar D, Mitchell P: **The impact of meaning and dimensionality on copying accuracy in individuals with autism.** *J Autism Dev Disord* 2006.
17. Mitchell P, Ropar D: **Visuo-spatial abilities in autism: a review.**
  - *Infant Child Dev* 2004, **13**:185-198.

Autistic drawings may reveal much about how low-level and high-level processing stages interact in vision; this reviews the relevant literature and current debates.
18. Mottron L, Belleville S: **Perspective production in a savant autistic draughtsman.** *Psychol Med* 1995, **25**:639-648.  
The perspective drawings of autistic draftsmen do not have a single vanishing point, showing that depth processing is local to regions within a scene.
19. Cavanagh P: **The artist as neuroscientist.** *Nature* 2005, **434**:301-307.  
This study introduces the concept of 'alternative physics' in visual art and that the violations of traditional physics in paintings do not appear aberrant.
20. Conway BR: **Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1).** *J Neurosci* 2001, **21**:2768-2783.
21. Lamme VA: **The neurophysiology of figure-ground segregation in primary visual cortex.** *J Neurosci* 1995, **15**:1605-1615.
22. Lee TS, Mumford D, Romero R, Lamme VA: **The role of the primary visual cortex in higher level vision.** *Vision Res* 1998, **38**:2429-2454.
23. Nakayama K, Shimojo S: **Toward a neural understanding of visual surface representation.** *Cold Spring Harb Symp Quant Biol* 1990, **55**:911-924.
24. Yazdanbakhsh A, Livingstone MS: **End stopping in V1 is sensitive to contrast.** *Nat Neurosci* 2006, **9**:697-702.
25. Zipser K, Lamme VA, Schiller PH: **Contextual modulation in primary visual cortex.** *J Neurosci* 1996, **16**:7376-7389.
26. Conway BR, Livingstone MS: **Spatial and temporal properties of cone signals in alert macaque primary visual cortex.** *J Neurosci* 2006, **26**:10826-10846.
27. Pack CC, Conway BR, Born RT, Livingstone MS: **Spatiotemporal structure of nonlinear subunits in macaque visual cortex.** *J Neurosci* 2006, **26**:893-907.
28. Jacobson J, Werner S: **Why cast shadows are expendable: Insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art.** *Perception* 2004, **33**:1369-1383.
29. Ostrovsky Y, Cavanagh P, Sinha P: **Perceiving illumination inconsistencies in scenes.** *Perception* 2005, **34**:1301-1314.
30. Kingdom FA, Beauce C, Hunter L: **Colour vision brings clarity to shadows.** *Perception* 2004, **33**:907-914.
31. Kingdom FA, Rangwala S, Hammamji K: **Chromatic properties of the colour-shading effect.** *Vision Res* 2005, **45**:1425-1437.
32. Kennedy JM, Juricevic I: **Foreshortening, convergence and drawings from a blind adult.** *Perception* 2006, **35**:847-851.
33. Hyer B: **Tonality.** In *New Grove Dictionary of Music and Musicians*, revised edn. Edited by Sadie aly. London: MacMillan; 2000.
34. Cook N: **The perception of large-scale tonal closure.** *Music Percept* 1987, **5**:197-205.
35. Freedman DJ, Assad JA: **Experience-dependent representation of visual categories in parietal cortex.** *Nature* 2006, **443**:85-88.
36. Kourtzi Z, DiCarlo JJ: **Learning and neural plasticity in visual object recognition.** *Curr Opin Neurobiol* 2006, **16**:152-158.
37. Gauthier I, Bukach C: **Should we reject the expertise hypothesis?** *Cognition* 2007, **103**:322-330.
38. McKone E, Kanwisher N, Duchaine BC: **Can generic expertise explain special processing for faces?** *Trends Cogn Sci* 2007, **11**:8-15.
39. Bukach CM, Gauthier I, Tarr MJ: **Beyond faces and modularity: the power of an expertise framework.** *Trends Cogn Sci* 2006, **10**:159-166.
40. Riesenhuber M, Jarudi I, Gilad S, Sinha P: **Face processing in humans is compatible with a simple shape-based model of vision.** *Proc Biol Sci* 2004, **271**(Suppl 6):S448-S450.
41. Connor CE, Brincat SL, Pasupathy A: **Transformation of shape information in the ventral pathway.** *Curr Opin Neurobiol* 2007, **17**:140-147.
42. Papaodysseus C, Fragoulis DK, Panagopoulos M, Panagopoulos T, Rousopoulos P, Exarhos M, Skembris A: **Determination of the method of construction of 1650 B.C. wall paintings.** *IEEE Trans Pattern Anal Mach Intell* 2006, **28**:1361-1371.
43. Halverson J: **The first pictures: perceptual foundations of Paleolithic art.** *Perception* 1992, **21**:389-404.
44. Sadr J, Jarudi I, Sinha P: **The role of eyebrows in face recognition.** *Perception* 2003, **32**:285-293.
45. Sinha P, Balas BJ, Ostrovsky Y, Russell R: **Face recognition by humans: nineteen results all computer vision researchers should know about.** *Proc IEEE* 2006, **94**:1948-1962.  
This study underscores the 'alternative physics' used by the brain to encode and recognize faces (see Cavanagh, 2005).
46. Costa M, Corazza L: **Aesthetic phenomena as supernormal stimuli: the case of eye, lip, and lower-face size and roundness in artistic portraits.** *Perception* 2006, **35**:229-246.
47. Sinha P, Poggio T: **I think I know that face.** *Nature* 1996, **384**:404.
48. Sinha P, Poggio T: **'United' we stand.** *Perception* 2002, **31**:133.
49. Tsao DY, Freiwald WA, Tootell RB, Livingstone MS: **A cortical region consisting entirely of face-selective cells.** *Science* 2006, **311**:670-674.