Colour Vision as a Scientific Problem

Arne Valberg

Our understanding of the environment largely based on interpretations of visual qualities, like contrast, form, and movement. No such property would be perceived were it not for the differences in colour (including white, grey and black). The qualities of colour – or qualia – belong to the realities of personal experience. One may, however, sometimes wonder about the link between one's conscious experience of colours, e.g. the redness of an apple, and the stimulation of the eye by electromagnetic radiation. Is there a causal relationship between a physical property of the apple's surface, the evoked neural activity and my perception of red? Is the spectral reflectance of the coloured surface the salient parameter, or is the activation of certain nerve cells the more essential? Or should we talk not so much about cause and effect, but rather, of complex scenes and neural states, of correlation and co-variation? Below we shall discuss some of these questions in the light of psychophysics and modern neuroscience (see also Valberg, 2001).

The quality of the visual experience 'red', or any other colour, cannot be identified with the spectral composition of light reaching the eye from the surface that gives rise to it (e.g. light reflected from the apple's surface). A surface with a certain spectral reflectance, can take on virtually any colour, depending less on the illumination and the reflected spectral distribution than on the surround conditions and the adaptation of the eye. This was well known in the last century and has more recently been emphasised by Edwin Land (1959, 1983). The colour is therefore not "caused" by the spectral composition of an isolated patch, or even the corresponding excitations of three cone types in the retina. However, this assumption is common, even among scientists.

The visual system of higher vertebrates is probably the most thoroughly investigated function of the brain, and colour perception is well suited to bring forth the various aspects of conscious experience. Today, the study of colour perception is closely tied to experimental psychophysics and neuroscience: even philosophical reflections about this issue are submitted to rather strong experimental constraints (Hardin, 1988; Thompson, 1995; se also the discussions of Palmer, 1999, and Saunders and van Brakel, 1997). Presumably, a neuroscientific program would search for collections of neurones whose activity is correlated with colour perception. One would record the activity of these neurones when colour stimuli change along the psychophysical dimensions *luminance, wavelength,* and *purity,* in the hope of finding links to the perceptive properties *lightness/brightness, hue,*

and *saturation* (*colour strength)*. Such attempts would give us useful overviews of the neural representations of different colour stimuli. This project would reach completion once we could find correlates between subjective properties of colour, the physical parameters of light projected on the retina, and the neural activities accompanying these (Churchland, 1994; Churchland and Sejnowski, 1992; Crick, 1994; Crick and Koch, 1995; Valberg et al., 1986, 1987). Some examples of such correlates will be described later in this paper.

Colour opponency as a problem within the neuroscientific program

In the years after Hermann von Helmholtz, colour perception was for a long time explained by light absorption in three types of cone photoreceptors, L, M, and S, in the human retina. The notion of the classical Young-Helmholtz trichromatic theory (that we now know to be wrong) was that three different classes of cone activated three basic colour processes or sensations: red (L-cones), green (M-cones), and blue (S-cones). All other colours were said to be due to the excitation of these primary sensations in different proportions (the perception of yellow would for instance result from equal stimulation of "green" and "red" processes, whereas white resulted from equal stimulation of all three).

This theory was strongly opposed by Ewald Hering's (1878; 1964) revolutionary claim that colour vision was based, not upon three primary sensations, but on six unique colour percepts (*Urfarben*) and their corresponding physiological processes. In Hering's theory, two pairs of chromatic unique hues, together with the achromatic pair, black and white, were associated with three pairs of antagonistic physiological processes. Unique red, for instance, was associated with the breaking down, or wearing out of a particular "visual substance" – and unique green with the building up, or restoration, of the same substance. Similar antagonistic processes in another substance gave rise to the yellow and blue unique hues. The opposite nature of these paired qualities, were thus associated with processes which mutually excluded each other. When the Swedish physiologist Gunnar Svaetichin (1956) made the first recordings of spectrally dependent positive and negative potentials in the retina of fish, he thought he had proven Hering to be right. Later, when Russell De Valois (1965) found spectral activation and inhibition of cells in the lateral geniculate nucleus (LGN) of the macaque monkey, these processes were looked upon as further confirmations of the opponent theory. It soon became common to use colour terms when referring to opponent cells as in the notations "red-ON cells", "green-OFF cells", etc. (Wiesel and Hubel, 1966), and many psychophysicists did not hesitate to relate the unique colour pairs directly to opponency in these cells (the "labelled lines hypothesis").

The analogy was carried even further so as to imply that each pair of unique colours could be identified with excitation and inhibition of one and the same type of opponent cell. If green corresponded to excitation, red would be the result of inhibition (Clark, 1993, p.127). These are understandable misinterpretations in view of Hering's postulate, and they will be discussed below. Even if Hering's elementary – or unique hues are today accepted as subjective references in phenomenal colour perception, it is necessary to emphasise that no correlates to them in the form of unitary physiological processes have been discovered. Consequently, in theories of colour vision, the unique hues are still without a simple neural representation, and their physiological origin remains a mystery.

In modern neuroscience, the general term opponency is now replaced by "cone-opponency" because it is the opposing cone inputs to a ganglion cell that divides the spectrum into excitatory and inhibitory regions: one class of cone activates the cell (+ sign) and another suppresses its activity (- sign). With three different L-, M-, and S-classes of cone, several two-opponent cone combinations are in principle possible (e.g. L – M, M – L, etc.), but not all of them are realised in nature (Valberg et al., 1987; Lee, 1996).

Thus opponency has at least two meanings: what we may call "cone-opponency" of neurones in the retina and LGN and "colour opponency" in perception[1]. The consensus is emerging that the mere existence of cone opponent cells cannot be regarded as a proof of Hering's theory, and that the use of colour terms in referring to the responses of such cells is misleading and a loose terminology (Abramov and Gordon, 1997). The idea of paired opponent processes in "the same physiological substrate" to explain unique hues needs revision. Opponent cells exist in many species and at all levels of neural processing in the primate; in the retina, in the geniculate nucleus, and in visual cortex. But there is, logically speaking, no necessary correspondence between opponent processes and unique hues, except for a striking analogy of polarity.

Several explanations are possible for why opponency is as universal as it seems to be. For instance, using psychophysical techniques, it has been demonstrated that there are opponencies in vision that serve other purposes than those postulated by Hering. Opponent "cardinal directions" in colour stimulus space have been discovered that are special with respect to colour discrimination and adaptation. These and related experiments have been summarised and discussed in Thompson (1995, pp.67–71) and in Mollon and Jordan (1997). In addition, theoretical considerations have demonstrated that opponency is a strategy that can be used by the nervous system to achieve an efficient transmission of information from the retina to the brain. By subtracting the response of one class of cone from that of another, the information that is common to both – and which is therefore redundant – is effectively removed (Buchsbaum and Gottschalk, 1983). We thus find opponency between cone classes in several channels. Cardinal directions and optimal transfer of information have little, if anything to do with Hering's unique hues. Although it is still believed that also colour perception relies on elementary opponent processes, it has been hard to track down the responsible mechanisms.

In what follows, we shall briefly analyse what is known about the responses of opponent cells to stimuli that have been scaled subjectively in hue and chroma (colour strength). This analysis will lead us to the following conjecture: a neural correlate for your seeing a particular colour (say unique red) has not yet been found, but covariations with more abstract aspects of colour space, like constant hue and constant chroma, have. This is an important result that may serve as a restriction upon future colour vision hypotheses.

We do not intend to answer questions like "*what* is unique red". This and other questions of the same sort (*what* is force, *what* is matter, or movement, for example) are generally not answered by natural science. We would be quite contented with an account for *how* a particular colour attribute is achieved under a certain viewing condition.

Unique hues

Six particularly simple colours are found in nature (da Vinci, 1906). They are the four chromatic unique, or elementary hues, yellow, red, blue, and green, together with the two achromatic colours white and black. These colours serve as six qualitative references in subjective or phenomenological colour space. The physical stimuli associated with these percepts may, however, vary from person to person. Unique yellow, for instance, can be determined with an extraordinary precision of less than 1 nm (nanometer) in the spectrum, although one person's selected wavelength may be found anywhere between 565 nm and 590 nm.

Unique yellow is characterised by it being "neither reddish nor greenish". It is thus determined purely subjectively by means of the two closest unique hues on the hue circle. Unique blue satisfies the same definition. This yellow-blue pair is called opponent because the two colours mutually exclude one another. No colour is normally seen as both yellow and blue at the same time, in the way that orange can be said to be perceptually composed of yellow and red. The same reasoning applies to the unique red-green pair. These definitions seem to have little to do with culture and language and with the related problems of colour categorisation (discussed in Saunders and van Brakel, 1997).

Following Hering's concept, a pure percepual arrangement of hue qualities can be derived using subjective scaling techniques. One example is the symmetrical hue circle of the Natural Colour System (NCS), where hues are scaled as proportions of the elementary hues yellow and red, red and blue, blue and green, and green and yellow. This systematic arrangement can for instance be visualised by a phenomenological hue circle consisting of four quadrants. A consequence of this arrangement is that the number of hue steps is much larger in the blue-red and blue-green sectors than in the yellow-red and yellow-green sectors. On a scale with forty equal hue steps around the whole circle, like in the Munsell system, the number of hue steps between unique blue and unique red, and between unique blue and unique green, is about 12. Between unique red and unique yellow, and between unique yellow and unique green, it is about 8.

The behaviour of a network of cone-opponent cells

As we have seen,the cone signals that are of importance for colour vision are mediated by neural interactions in the retina among cone-opponent retinal ganglion cells. Ganglion cells in the retina project to opponent cells in the lateral geniculate nucleus (LGN), and from there to the primary visual cortex. Opponent cells are particularly sensitive to differences in the activation of cone classes, and consequently they are coding for wavelength differences. Let us imagine

that we are looking at a red traffic light. An opponent, 'L-M' cell would be strongly activated by this long-wavelength light. Its signal is proportional to the difference, $N = V_L - V_M$, of the responses, V, of L- and M-cones. Because L-cones are excited by long-wavelength light and M-cones by mid-spectral light, the result of looking at the red traffic light is more excitation in the L-cones than in the M-cones, and the difference $N = V_L - V_M$ becomes positive. As soon as the eye is exposed to the red light, this 'L-M' cell will transmit a train of electrical impulses up the optic nerve to the lateral geniculate nucleus, the next station in the brain. When the light shifts to green, this cell immediately stops firing because of the strong inhibition from M-cones, giving a negative value of the difference. The green light activates 'M-L' cells, the opposite cell type of 'L-M'. In the case of achromatic white light (as well as for the yellow traffic light) L- and M-cones are excited about equally, cancelling each other's input to an opponent cell, and the responses of both 'L-M' and 'M-L' cells to white will therefore be rather small[2].

In *Table I* we find the colour combinations that optimally stimulate or inhibit these cells under a normalised, neutral viewing condition (Valberg et al., 1986; Valberg 1998):

Under other experimental conditions, with a different white adaptation or with a different surround, unique red would be associated with another spectral distribution, and one would therefore expect another ratio of responses between the opponent mechanisms. Further experiments are needed in order to answer the question as to what extent these opponent cells' responses are linked solely to the physical properties of the stimulus, or if the colour actually perceived in a particular viewing situation is also important. Consequently, we have not yet discovered the neural correlate of your seeing red (or any other colour). The ratio of cell responses is only a covariation, in a normalised viewing situation, with your perceiving a constant hue. The former correlate, the so-called "redness" of red, must be looked for in higher-level brain activities.

To summarise, a neural representation can account for two different attributes, namely constancy of hue and constant chroma. A particular ratio of responses of two neighbouring cell types correspond neatly with a constant hue, whereas response magnitude relates (in a yet to be determined way) to colour strength or chroma. This behaviour clearly resembles vector coding. However, this does not mean that one knows which hue is associated with a particular ratio of relative responses of opponent cells. To establish such a relationship, additional information is needed as to how the experimental situation, like simultaneous contrast, adaptation, stimulus size, temporal aspects, etc., influence cell responses.

Conclusion

There may be several stages of transformations of neural signals before "the level of perception" is reached – if, indeed, such a distinct higher level really exists, except as an abstraction of our minds (Dennett, 1996). Only a combination of psychophysical and neurophysiological methods, together with quantitative modelling, will bring us closer to an understanding of the complex processes involved, and draw attention to the limits of current psychophysical linking hypotheses (Lee, 1991).

It may not be too surprising that it is easier to find a geometrical representation of *equalities* (matches) and *differences*

Table I

Opponent cells	Colours in a normalised viewing condition
L – M	activated by bluish red lights (Munsell 5RP) and inhibited by greenish lights
M – L	activated by bluish green lights (Munsell 10G) and inhibited by reddish lights
M – S	activated by greenish yellow lights (Munsell 2.5GY) and inhibited by bluish lights
S – L	activated by reddish blue lights (Munsell 2.5P) and inhibited by yellowish lights

(discrimination) than for *perceptual qualities* (contents of conscious experience). Natural science can describe the physical and physiological conditions under which we see colours, by referring to biophysical states of a complex system. This is probably all one can expect from a scientific explanation. What we can strive for is to give a quantitative account of the physical and physiological conditions that give rise to the *same* percepts; i.e. where an attribute of a colour stimulus in one condition matches the same attribute in another condition. Then it does not matter which colour it is, or if it is of unique- or binary hue.

It is clear that several opponent channels are sharing a response that somehow contributes to a constant percept, and that a low-level single channel thus transmits information about more than one property. Such multiplexing of functionally relevant information in one cell type, or in one channel, means that cells at the lower levels cannot distinguish between properties like size, contrast, intensity, form, or colour. According to a bottom-up view such properties must first be separated so that they at a later stage can become entities of conscious perception.

We conclude that we cannot come much closer to *what* colour qualities and unique colours are, but we can use a scientific language to point to the circumstances that are required for their appearance – and to investigate their evolution. Thus, our subjective colour space, with its unique colour references, is mapped onto a stimulus space in an ever changing relationship, depending on adaptation, surround conditions, etc.. Neuroscience makes it possible to investigate some of these relationships experimentally.

Acknowledgements
I thank Inger Rudvin for helpful comments. This work has been supported by a grant from Thonning Owesen foundation.

Notes

1. Obviously, the "redness of red" does not have a simple physical correlate. Even if we can specify the many different physical and non-physical conditions that normally lead to a red percept, we cannot explain why we experience redness and not another quality instead. Neither is the perception of red correlated with the physiological stage of receptor excitations as Land's demonstrations (1959, 1983) so vividly reminded us. As far as we know, at the next several levels of neural firing there is still non-correspondence.
2. There are two subtypes of each of the 'L – M'- and of the 'M – L' cells. The main difference between the two subtypes is the relative weights of excitation (+) and inhibition (-). One group has little inhibition and therefore responds well to bright lights, whereas the other has strong inhibition in bright lights and thus exhibits a weaker response, but they display clear opponency for darker colour stimuli (Valberg et al., 1986; Valberg, 1998). In addition, the retinal and geniculate cells have a particular spatial arrangement of excitatory and inhibitory cones within their receptive field that cause them to differentiate between not only particular colours, but also stimulus size, changes of light intensity, luminance contrast, etc. This multiplexing at lower levels of processing must be sorted out at later levels to arrive at useful information. Since perceiving all of an object's qualitative dimensions requires a conscious mind, one may therefore argue that the multiplexed information must be sorted out as properties before they reach consciousness. These properties together constitute the entity that one finally perceives as the object.

Arne Valberg, Professor
Institute of Physics, Dep. of Biophysics
Norwegian University
of Science and Technology
Trondheim, Norway
Arne.Valberg@phys.ntnu.no

References

ABRAMOV, I. & GORDON, J. 1997. Constraining color categories: The problem of the baby and the bath water. In *Behavioral and Brain Sciences* **20**, 179–180.

BUCHSBAUM, G. & GOTTSCHALK, A. 1983. *Trichromacy, opponent-colours coding and optimum colour information transmission in the retina*. Proceedings of the Royal Society London B**220**, 89–113.

CHURCHLAND, P.M. 1994. *The Engine of Reason, the Seat of the Soul.* The MIT Press, Cambridge, Mass.

CHURCHLAND, P.S. & SEJNOWSKI, T.J. 1992. *The Computational Brain.* MIT Press, Cambridge, Mass.

CLARK, A. 1993. *Sensory qualities.* Oxford University Press, Oxford.

CRICK, F. 1994. *The Astonishing Hypothesis. The Scientific Search for the Soul.* Charles Scriber's Sons, New York.

CRICK, F. & KOCH, C. 1995. Why neuroscience may be able to explain consciousness. Comment to Chalmers (1995b). *Scientific American*, December 1995,66–67.

DA VINCI, LEONARDO. 1906. *A Treatise on Painting*. Translated to English by J.F. Rigaud. G. Bell, London.

DENNETT, D.C. 1996. The myth of double transduction. Tucson-II conference on consciousness. Abstract

DE VALOIS, R.L. 1965. Analysis and coding of color vision in the primate visual system. Cold Spring Harbor Symposia on Quantitative Biology **30**, 567–579.

FARRELL, B.A. 1962. Experience. In *The Philosophy of the Mind*, edited by V.C. Chappell. Prentice-Hall, N.J., pp. 23–48.

HARDIN, C.L. 1988. *Color for Philosophers.* Hackett, Indianapolis.

HERING, E. 1964/1878. *Outlines of a Theory of the Light Sense.* Translated from German to English by L.M. Hurvich and D. Jameson. Harvard University Press. Orginal title: *"Zur Lehre vom Lichtsinne"* (Wien, Gerolds Sohn, 1878).

LAND, E.H. 1959. Color vision and the natural image. Part I and II. Proceedings of the National Academy of Sciences of the U.S.A. **45**, 115–129 and 636–644.

LAND, E.H. 1983. Recent advances in retinex theory and some implications for cortical computations: Color vision and the natural image. Proceedings of the National Academy of Sciences of the U.S.A. **80**, 5163–5169.

LEE, B.B. 1996. Receptive field structure in the primate retina. Vision Research **36**, 631–644.

LEE, B.B. 1991. On the relation between cellular sensitivity and psychophysical detection. In *From Pigments to Perception; Advances in Understanding the Visual Process.* (Ed. A. Valberg and B.B. Lee). Plenum Press, London, pp. 105–116.

MOLLON, J.D. & JORDAN, G. 1997. On the nature of unique hues. In *John Dalton's Colour Vision Legacy* (Ed. C. Dickinson, F. Murrey, and D. Carden). Taylor & Francis, 381–392.

PALMER, E.S. (1999). Color, consciousness, and the isomorphism constraint. (With open peer commentary). Behavioral and Brain Sciences **22**, 923–989.

SAUNDERS, B.A.C. AND VAN BRAKEL, J. (1997). Are there nontrivial constraints on colour catergorization? (With open peer commentary). Behavioral and Brain Sciences **20**, 167–228.

SVAETICHIN, G. 1956. Spectral response curves from single cones. Acta Physiologica Scandinavia **39**, Suppl. 134, 17–46.

THOMPSON, E. 1995. *Colour Vision*. Routledge, London.

VALBERG, A. & SEIM, T. 1983. Chromatic induction: responses of neurophysiological double opponent units? Biological Cybernetics **46**, 149–158.

VALBERG, A., SEIM, T., LEE, B.B., & TRYTI, J. 1986. Reconstruction of equidistant color space from responses of visual neurones of macaques. Journal of the Optical Society of America A**3**, 1726–1734.

VALBERG, A., LEE, B.B., & TRYTI, J. 1987. Simulation of responses of spectrally-opponent neurones in the macaque lateral geniculate nucleus to chromatic and achromatic light stimuli. Vision Research, 27, 867–882.

VALBERG, A. 1998. *LYS SYN FARGE.* Tapir, Trondheim.

VALBERG, A. 2001. Unique hues: An old problem for a new generation. Vision Research. In press.

WIESEL, T. & HUBEL, D.H. 1966. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. J. Neurophysiology 29, 1115–1156.