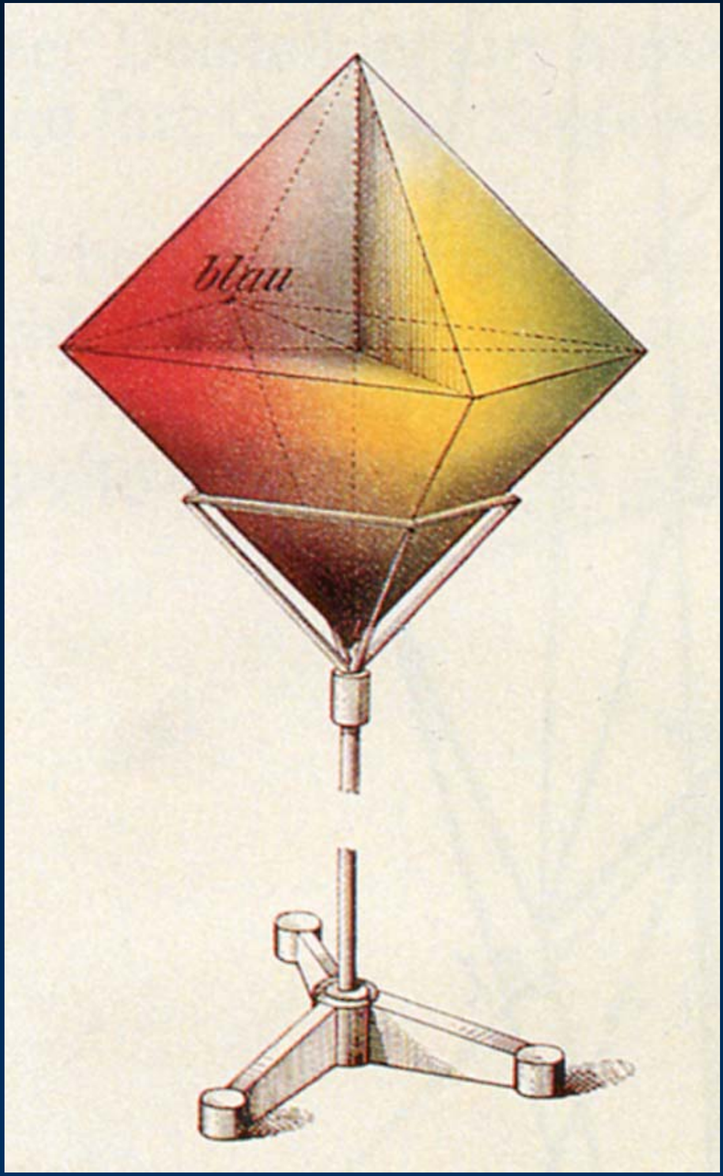
Color Spaces

Rolf G. Kuehni



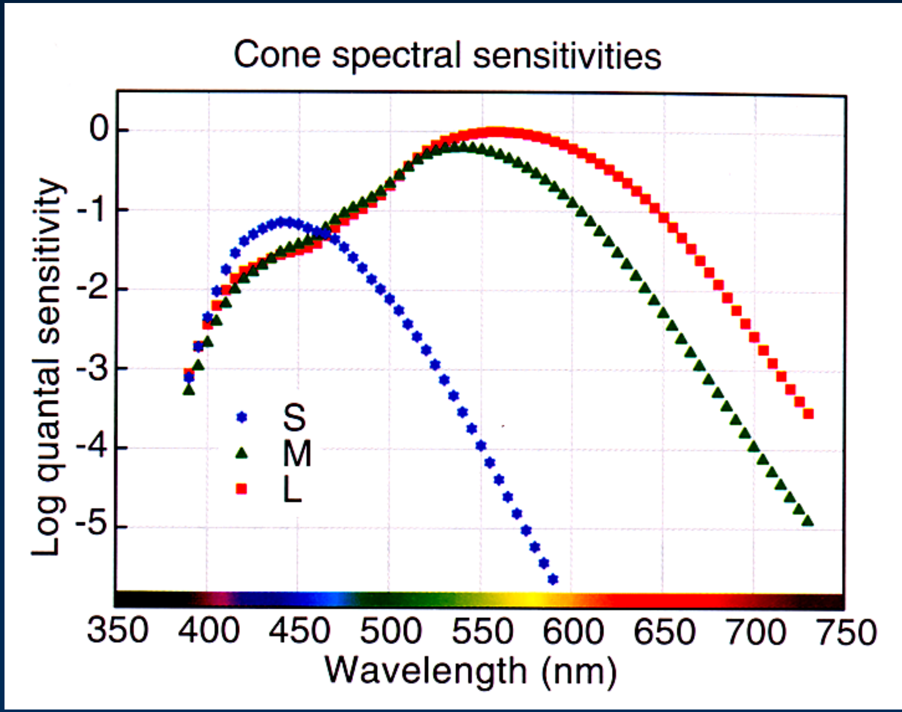
A Color stimulus spaces

B Colorant and light source mixture spaces

C Empirical perceptual spaces
- Hering type spaces
- isotropic spaces

D Stimulus spaces fitted to perceptual spaces

• Normal human color vision is mediated by three types of receptor cells in the retina called cones with different but overlapping spectral sensitivities.



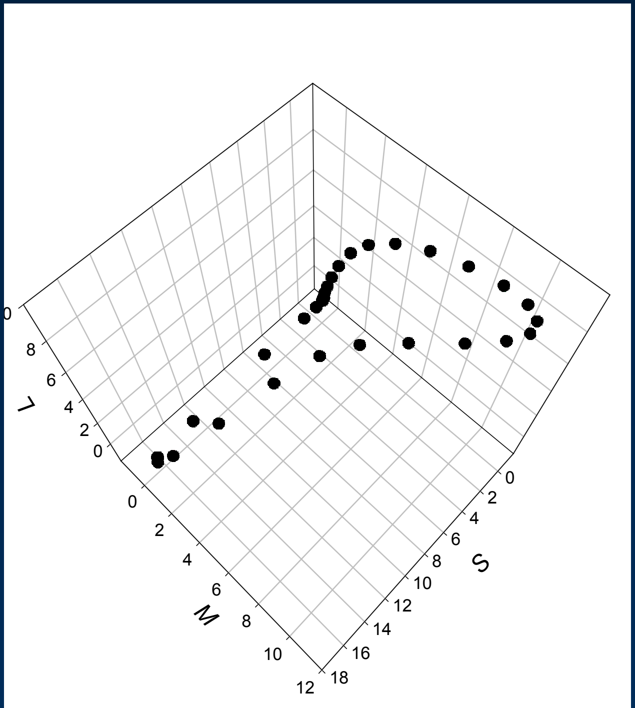
Sharpe et al. 1999

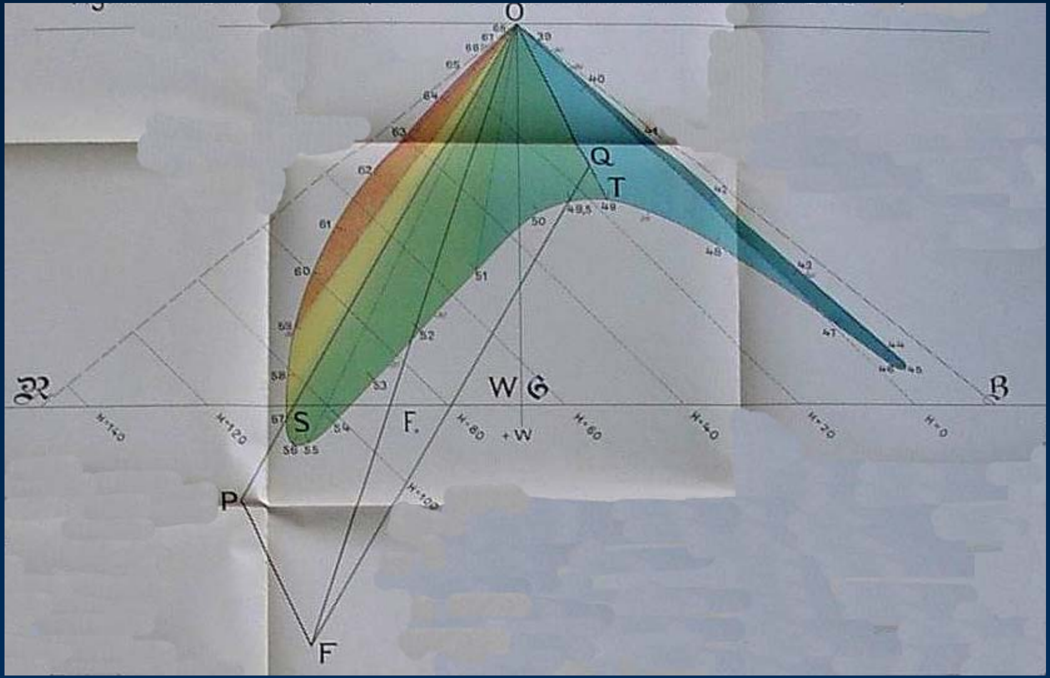
They are usually identified with L, M, and S for long, medium, and short-wave sensitivity.

They are the interface translating electromagnetic energy (light) from the outside world as input for the visual system.

• Different animal species have different numbers of light sensors and different spectral sensitivity ranges.

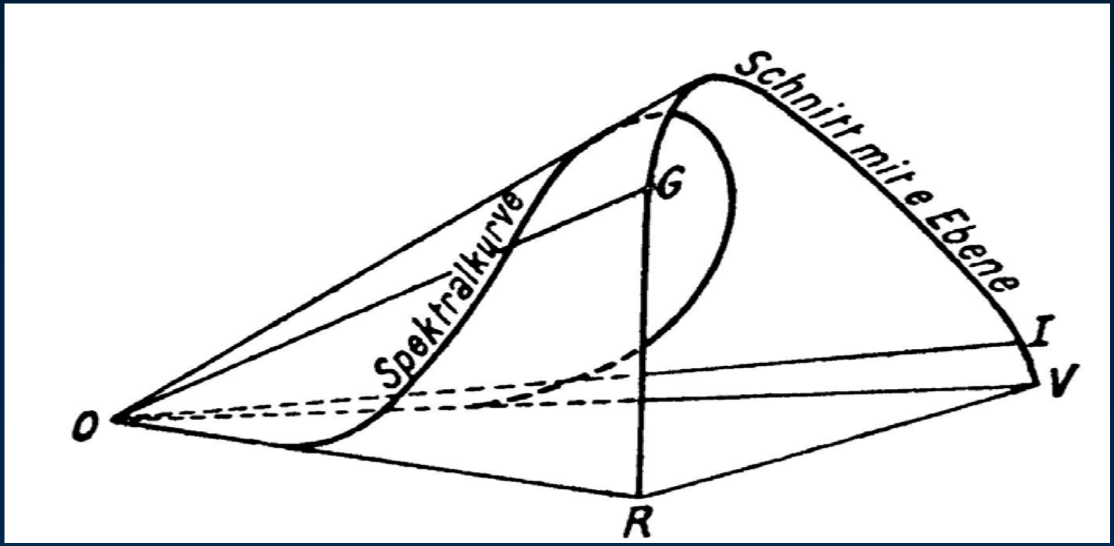
• A 3-dimensional color space is implicit in a system with three receptors (three cone types).
• When spectral lights are ‘filtered’ through the three cone functions their locus in the corresponding LMS cone space forms a curved line.





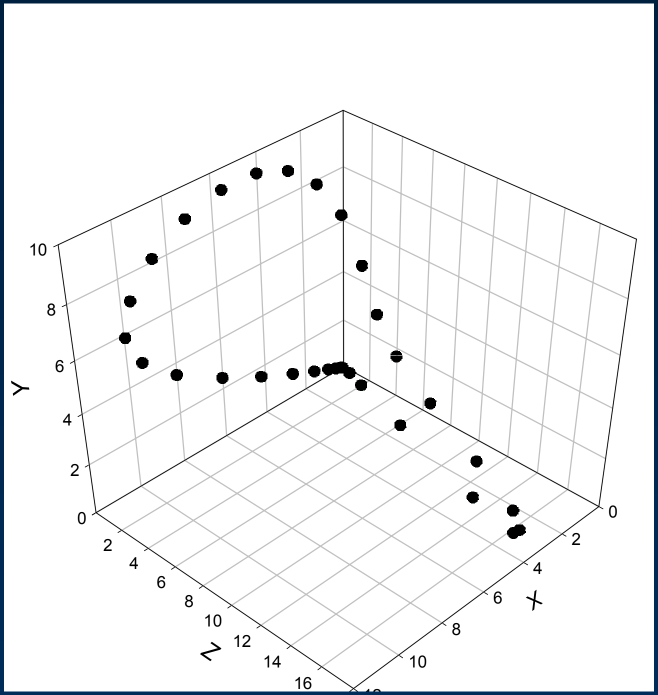
Pilgrim, 1901

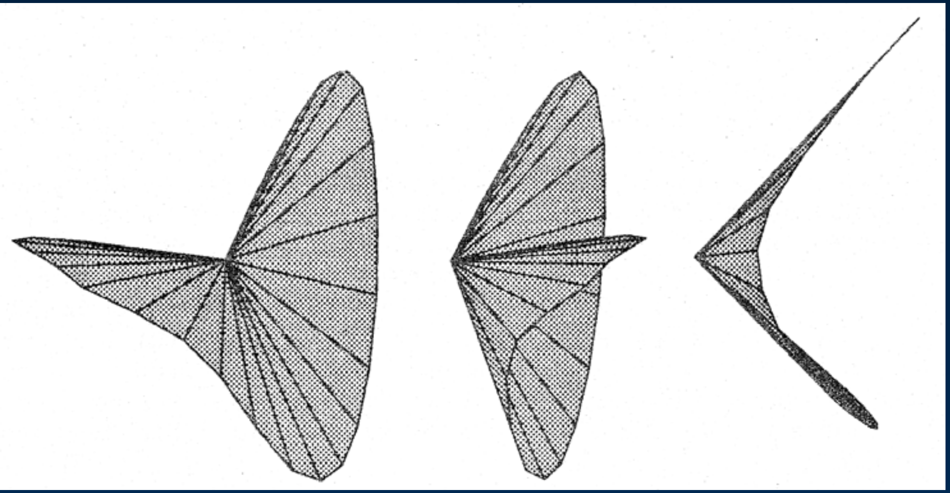
All object color stimuli have to fall within the ‘butterfly’-shaped shell formed by the spectral vectors.



Schrödinger (using red, green and violet primaries) named the resulting envelope "spectrum bag." All color stimuli have to fall inside.

• The axes of the space can be are linearly transformed so that one of them is in approximate agreement with brightness perception, as in the CIE tristimulus space.

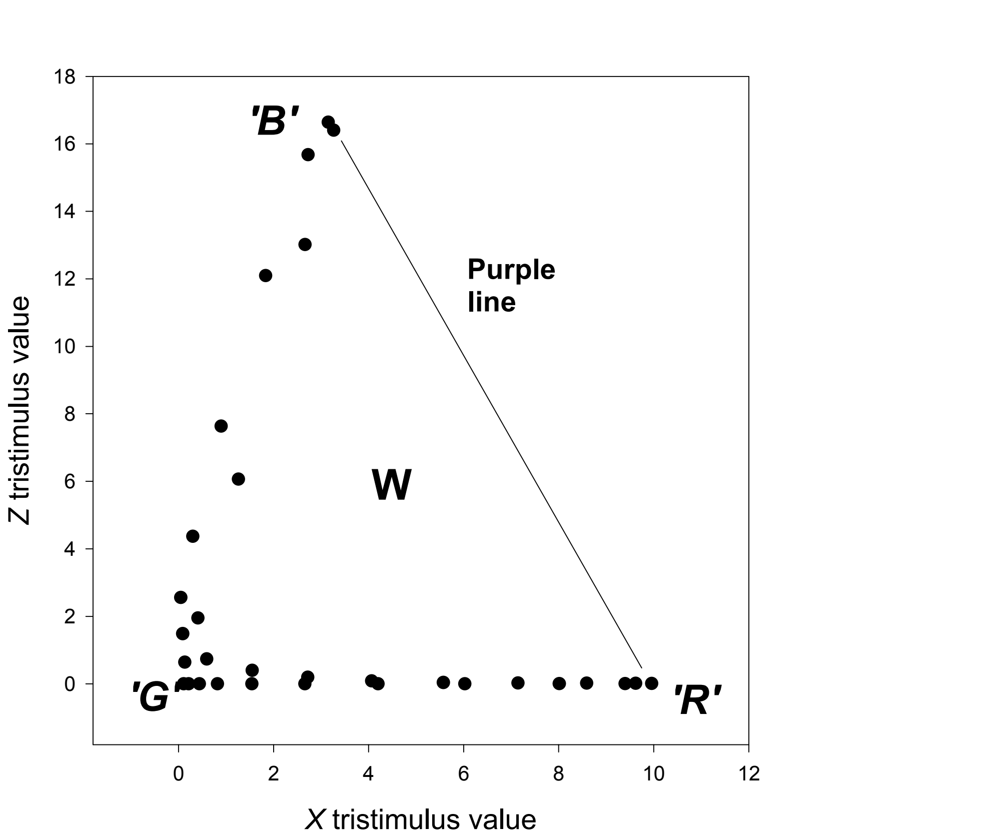




Three views of the surface of spectral stimuli vectors from different angles.

After Cohen

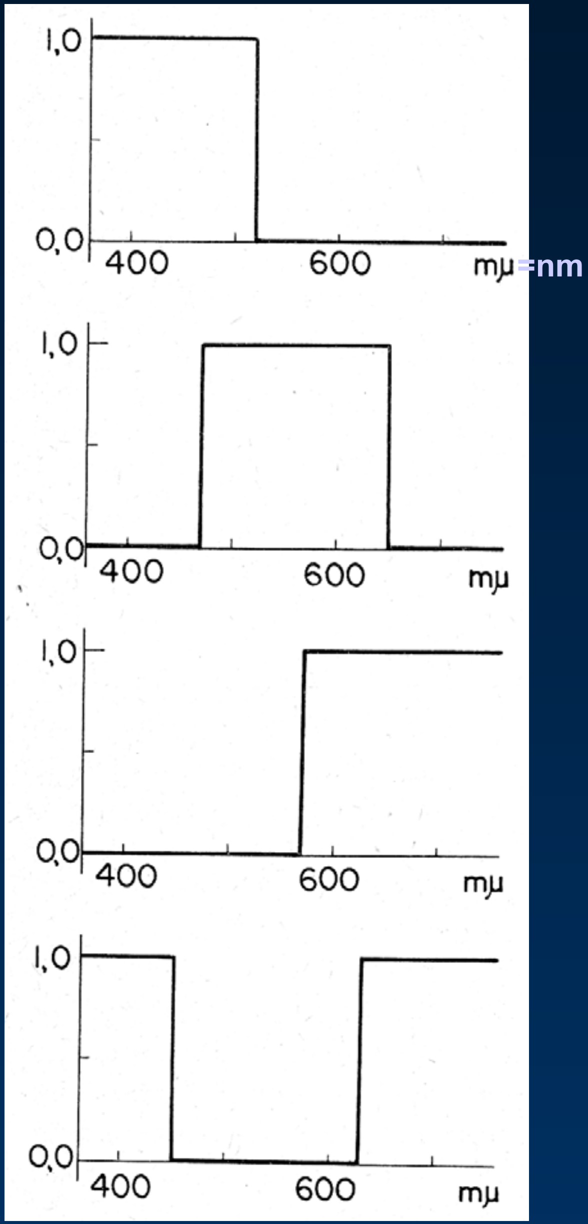
When projecting the endpoints of the spectral vectors onto the X, Z plane (containing information ‘other than brightness’) a triangular diagram is obtained.



It is the basis of Maxwell’s triangle.

The ‘purple line’ on which mixtures of spectral ‘R’ and ‘B’ are located represents the border in that direction for real stimuli.

***W* represents the equal energy or ‘white’ point of the diagram.**



Of interest are *optimal object color stimuli*:
These are the stimuli that at a given lightness level are the *farthest away* from the equal energy point. They have been defined in terms of reflectance functions. There are two fundamental types:

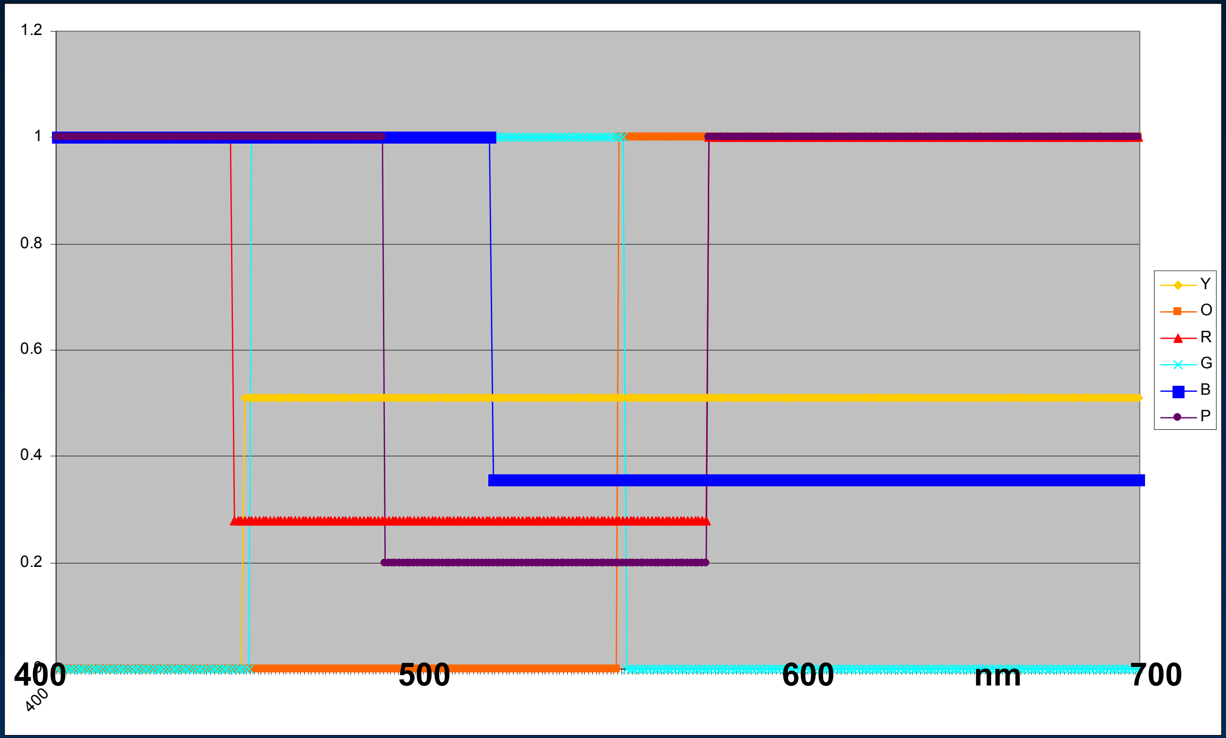
***single* transition: top and third Figure,**
***double* transition: second and fourth figure. Transition wavelengths can *vary across* the spectrum. At different levels of lightness the height of the loop varies.**

Surface of optimal object colors

- By systematically moving the transition wavelengths in all four cases through the spectrum and adjusting the height for brightness the reflectance spectra of all possible optimal object color stimuli can be generated. The assumed light source has equal energy across the spectrum.

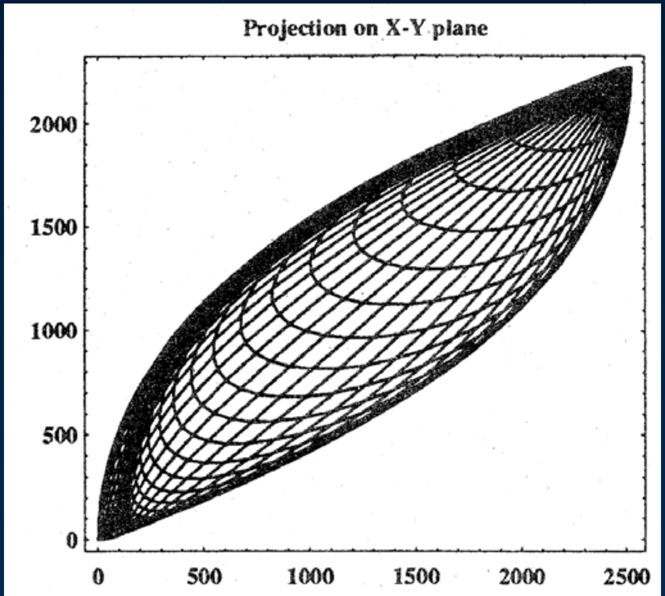
- When these reflectance spectra are projected into a tristimulus space they form the surface of optimal object colors within that space.

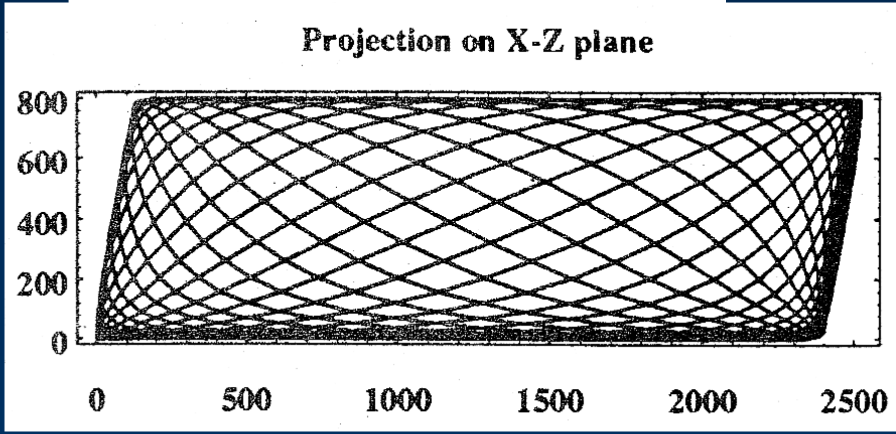
- Real object spectra differ from the optimal ones by being curved and by having reflectance values higher at the low end and lower at the high end than optimals.



Reflectance functions of six optimal object colors at Y = 50.

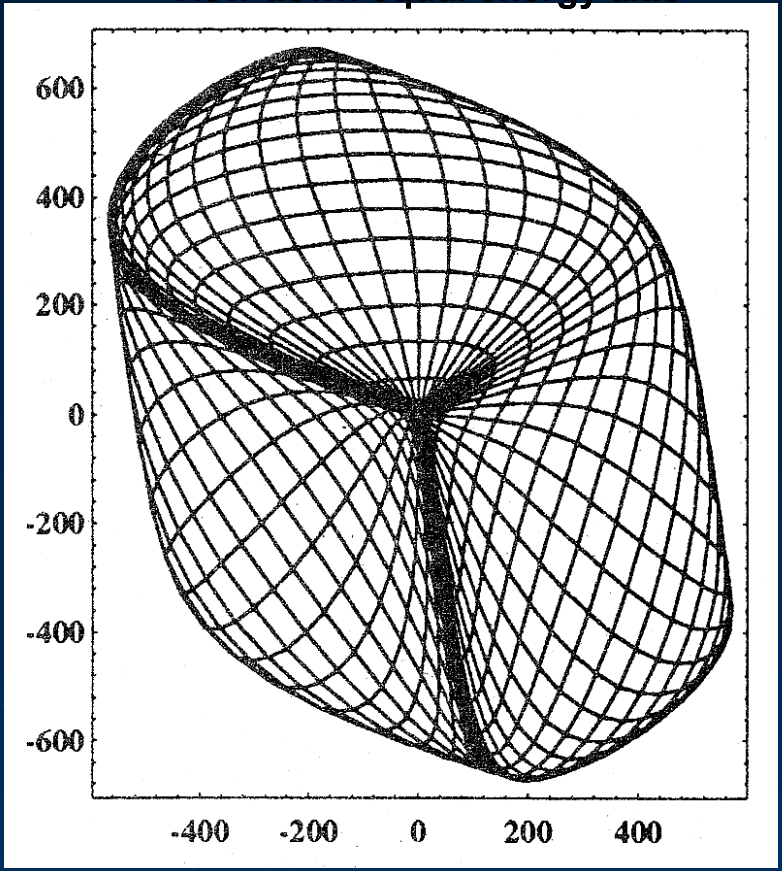
Surface of optimal object color stimuli in the CIE tristimulus space





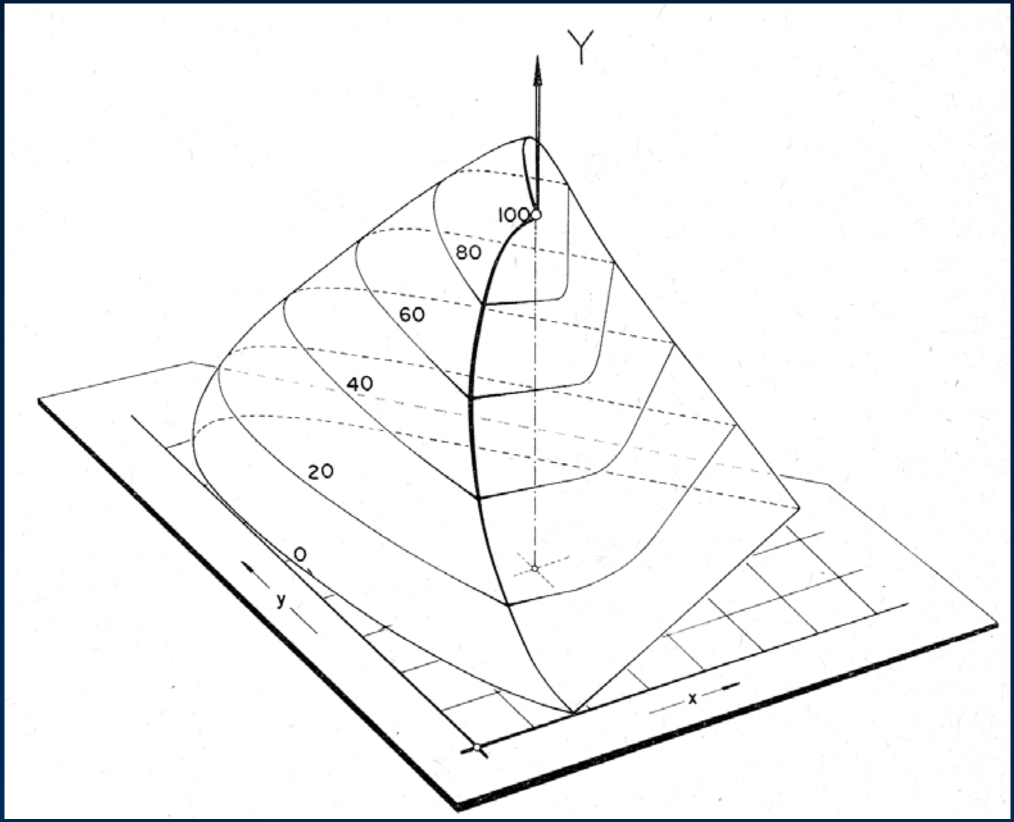
Tristimulus values

View down equal energy axis



Figures by Koenderink

There are various versions of color stimulus spaces based on the CIE colorimetric system. They all contain the same information if in different reference frames. Some have been encountered already. Another well-known one is the *xyY* space. It s basis plane is the CIE chromaticity diagram. The example shows the object color limits as viewed in standard daylight in that space.



Wyszecki, 1960

Another cone-based stimulus space

A different kind of cone stimulus space has its axes based on cone input into brightness and opponent cells as experimentally determined in the lateral geniculate nuclei (in mid-brain) of macaque by Derrington, Krauskopf and Lennie (1984).

The axes are defined as:

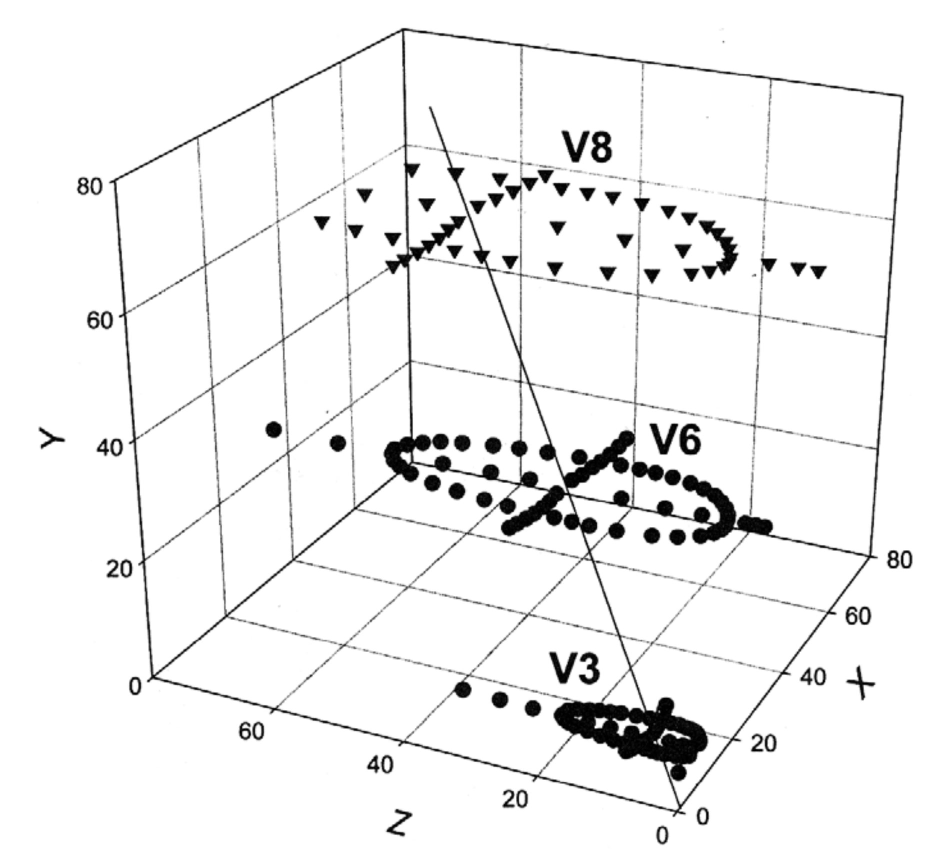
- 'Brightness': L+M
- 'Yellow-blue': L + M – S
- 'Red-green': L-M

The hue designations have only very general categorical meaning.

This space is widely used by color vision scientists to represent experimental data in a form they consider more meaningful than if presented in the CIE system.

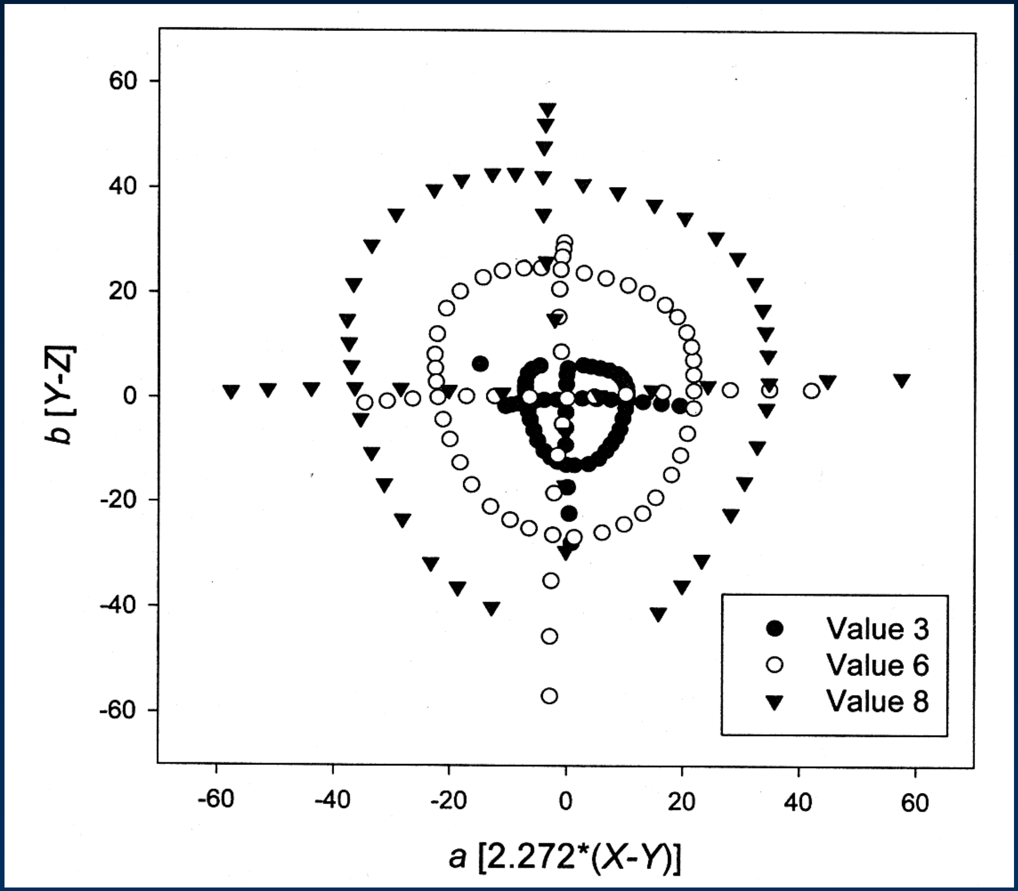
Object color stimuli

- All real object color stimuli are located inside the optimal object color stimulus surface; for example the Munsell color stimuli in CIE tristimulus space.



Munsell constant chroma hue circles, with selected color stimuli varying in chroma, at 3 levels of value in the *XYZ* tristimulus space. Contours are elongated along the *Z* axis.

• Equal energy points at different levels can be normalized.



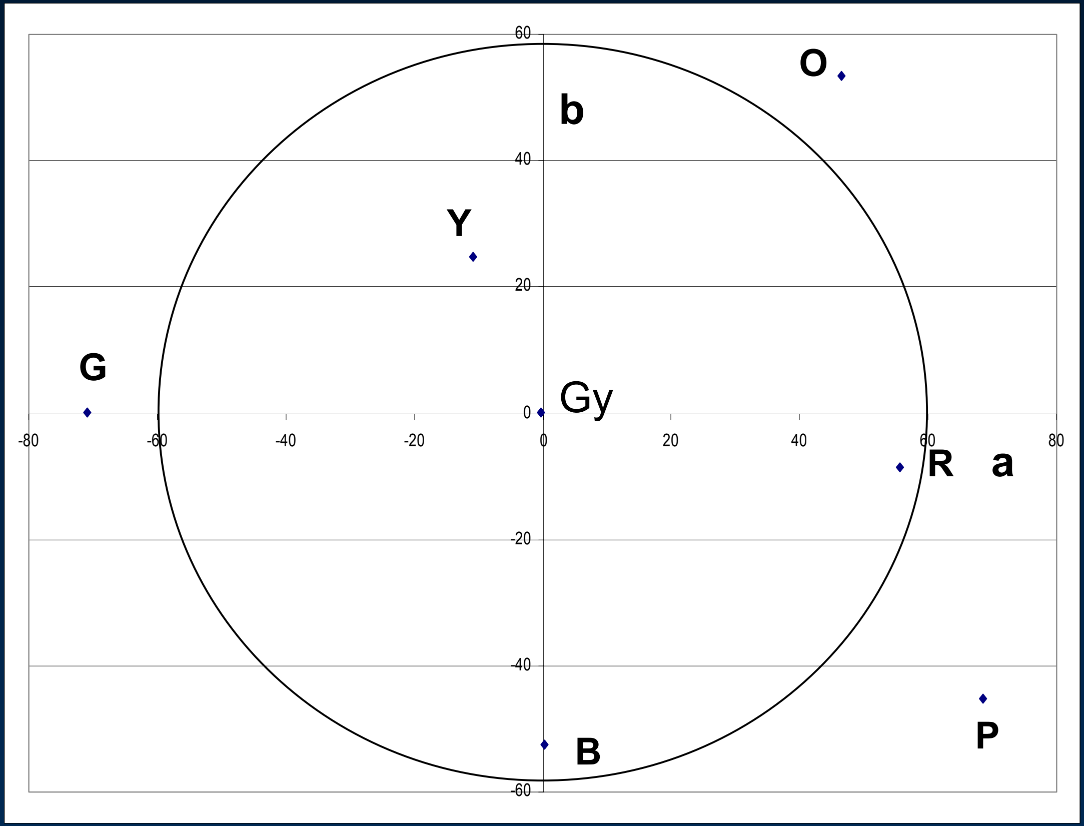
Projection of the three Munsell hue circles of the previous figure onto the normalized Cartesian basis plane replacing the X, Z plane.

The *a* axis has been weighted to normalize it with the *b* axis.

The first normalization is achieved by subtracting the value of the brightness dimension (vertical) from the values representing the two base plane dimensions.

As a result of this mathematical operation the base plane is converted to a standard Cartesian diagram with positive and negative values and the ‘spindle’ of the optimal object color surface in CIE tristimulus space is righted so that at all lightness values equal energy stimuli have coordinate values of 0, 0 in the basis plane.

Neurophysiologically, this operation is accomplished (in principle) by opponent color cells, as seen in the DKL space. It is an economical implementation of the mathematical operation.



Location of the six optimal object color stimuli of page 10 and the achromatic stimulus in the balanced *a, b* opponent color diagram (Y = 50). Color stimuli of constant chroma 60 fall on the circle. The hue circle is the result of dimension reduction and normalization.

- Steps taken so far:

  - filtering of spectral and broadband stimuli through cone functions
  - rotating axes so that one approximately corresponds to brightness perception
  - normalizing the equal energy point to be in the center of a Cartesian diagram regardless of the value of the brightness dimension; normalizing the two axes of the diagram.

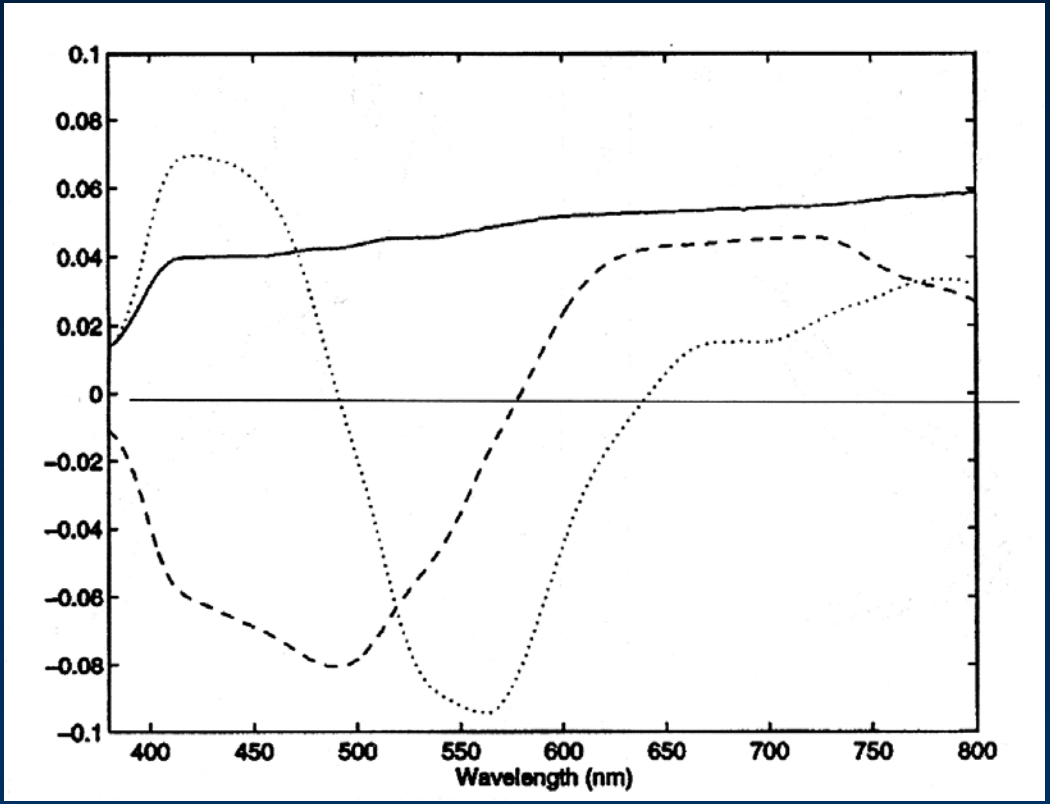
The result, in case of Munsell stimuli, is in ordinal agreement with their corresponding perceptual organization.

Optimized dimension reduction

The problem can be approached from another perspective:
The assumption is that the human visual system attempts to extract a high level of information from the spectral stimuli entering the eye.

- Such stimuli have many dimensions, at 1 nanometer spectral interval some 350 (730 nm – 380 nm). Filtering the stimuli through three ‘filters’ represented by the cone cell sensitivity functions reduces the dimensionality from 350 to 3. Some information is lost in the process, that is, the exact form of the spectral information cannot be recreated. But with *L, M, S* filters for a reduced range of 404-658 nm (main activity range of the 3 cone types) on average 86% of the variation in Munsell stimuli can. Improved levels of accuracy are obtained from reduction filters mathematically optimized against the stimuli in a set.

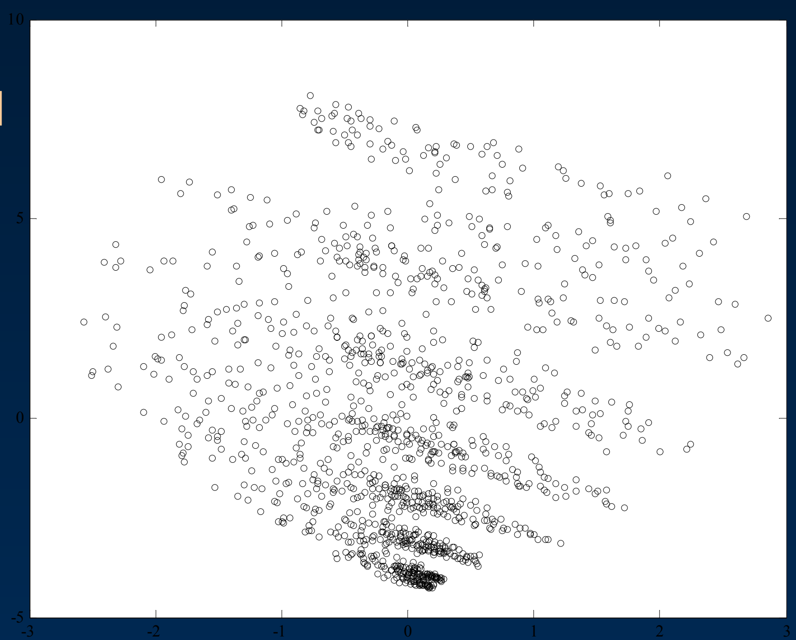
• There are different mathematical methods for dimensionality reduction, one of them is principal component analysis (PCA). The three most important PCA functions (Eigenvectors)resulting from analysis of the reflectance spectra of all Munsell color chips are shown:

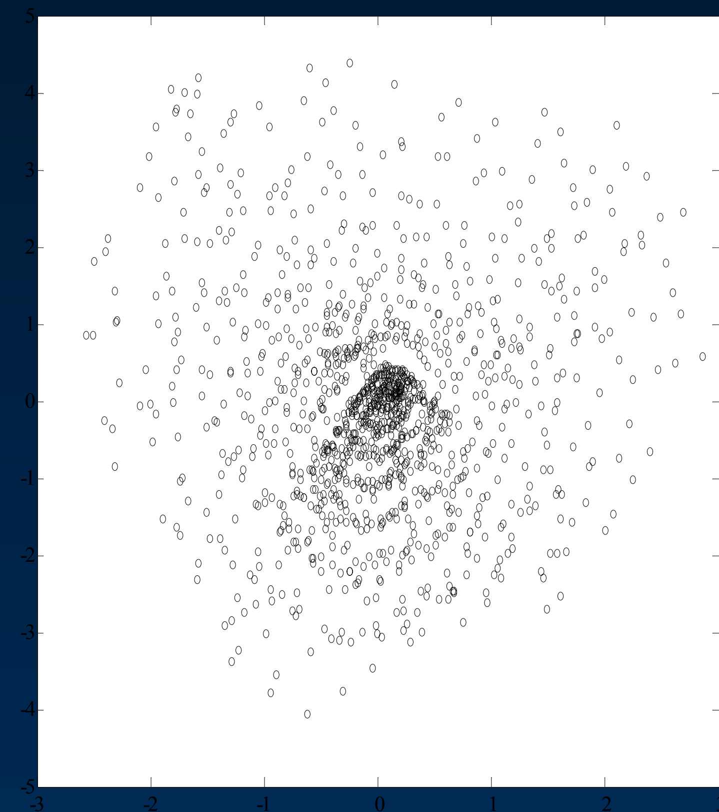


Lenz et al, 1996

The nearly flat function explains some 80% of the variation in the Munsell reflectance functions. It has a crude resemblance to a brightness function. The other two PCA functions resemble opponent color functions. (The set consists of selected data, it does not contain any metamers.) All three functions explain ca. 95% of the variation.

Color stimulus spac

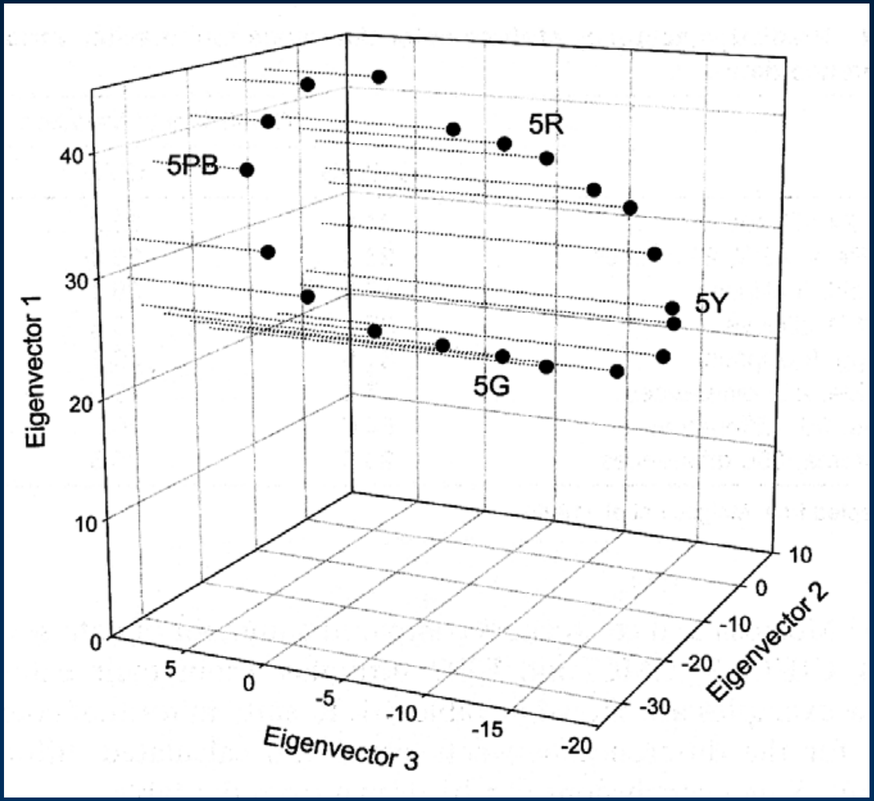




Left: 1310 Munsell reflectances in the EV2, EV1 plane
Right: the same reflectances in the EV2, EV3 plane

PCA places these spectra onto roughly circular loci (right) separated according to the average elevation of the curve in the third dimension (left).

• The reflectance functions of twenty Munsell color chips forming a hue circle at chroma 8 and value 6 are ordered in the PCA space of the first three Eigenvectors as shown here:



They are placed in correct ordinal order, compared to perceptual results, around the hue circle, but Eigenvector 1 does not represent Munsell value (lightness) closely.

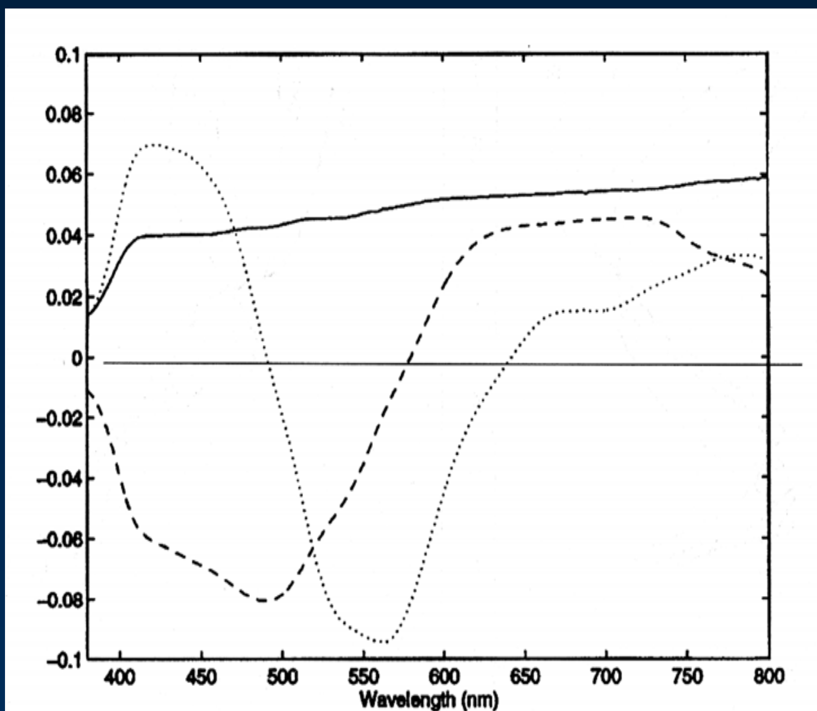
- The implication is that the hue circle is the natural outcome of dimension reduction of spectral functions having any possible spectral form to three dimensions.

- A circular contour is generated as long as at least two dimension reduction functions are themselves curved and overlapping. It is the result of mathematics alone.

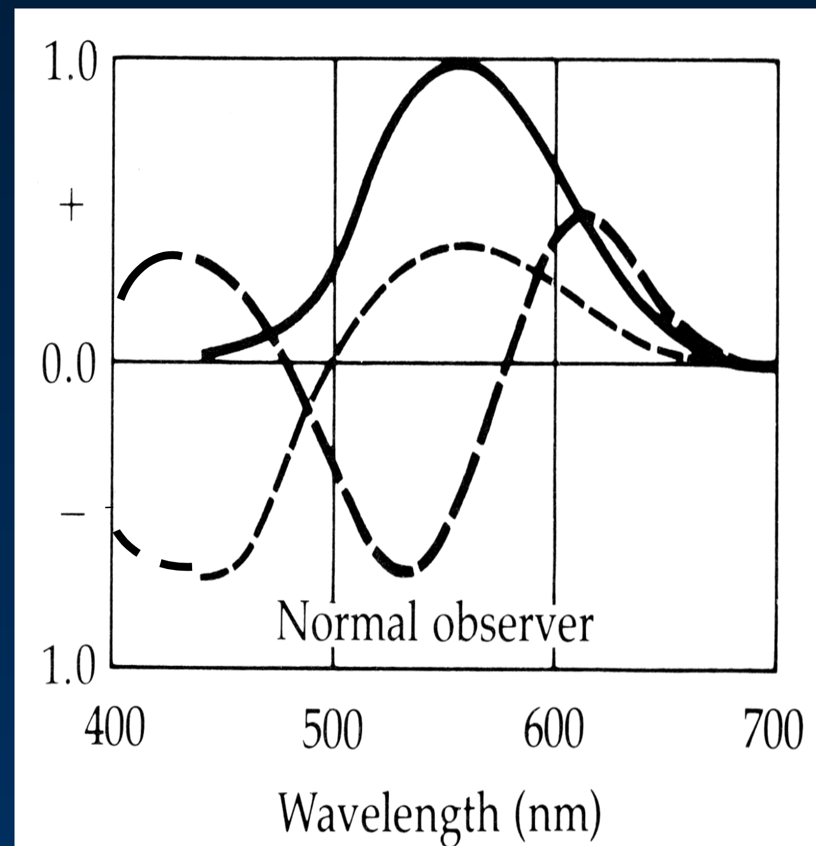
- Dimension reduction with three all positive functions places all object color reflectance functions into a spindle-like solid in the three-dimensional space.

- If two of the dimension reduction functions have positive and negative values (as the optimized PCA functions) the spindle is righted, resulting in normalized plots.

Comparison of PCA and cone based dimension reduction functions:



CIE standard observer luminance and opponent color functions



- The first three PCA Eigenvectors allow recovery of the spectral information of Munsell chips in the wavelength range from 380-730 nm at the 95% level. Five to six Eigenvectors are required to recover the spectral information at the 99+% level of accuracy. With *L, M, S* cone functions recovery is at the 66% level. With *L* and *S* alone it is 63%. In the reduced wavelength range of 404-658 nm *L, M, S* recovery rises to 86%.
- We can conclude:
  - Filtering with cone 'filters' is less than optimal.
  - Both kinds of filter place spectral functions of a large number of different object color stimuli in roughly circular form in ordinal agreement with hue perceptions from the same stimuli (under 'standard' conditions).
  - The circular order of hues from spectral and extra-spectral stimuli is an implicit result of dimensionality reduction by filtering with three filters (in the human case the cones).

We can conclude that a likely ‘goal’ of evolution was to have, within biological limitations, efficient recovery of the original spectral information (there are many complications however).

• Dimension reduction by filtering automatically results in an important problem. Certain classes of different stimuli cannot be distinguished by the system. In color science they have been called metamers (Ostwald). No examples of important natural metamers are known.

• Metamerism plays a significant *role* in the modern human world: it makes possible color photography, printing by half-tone and other methods, TV and video display, etc. It also causes large technological problems in color reproduction and control.

Relationship between spectral stimuli and color perceptions

This is an entirely different matter:

- its limitations are well known: due to
  - metamerism
  - contrast and adaptation effects
  - effects of quality of light on object appearance
  - variability in individual color matching functions
  - variability in selection of unique hue stimuli
  - definition of components of visual fields

  etc.

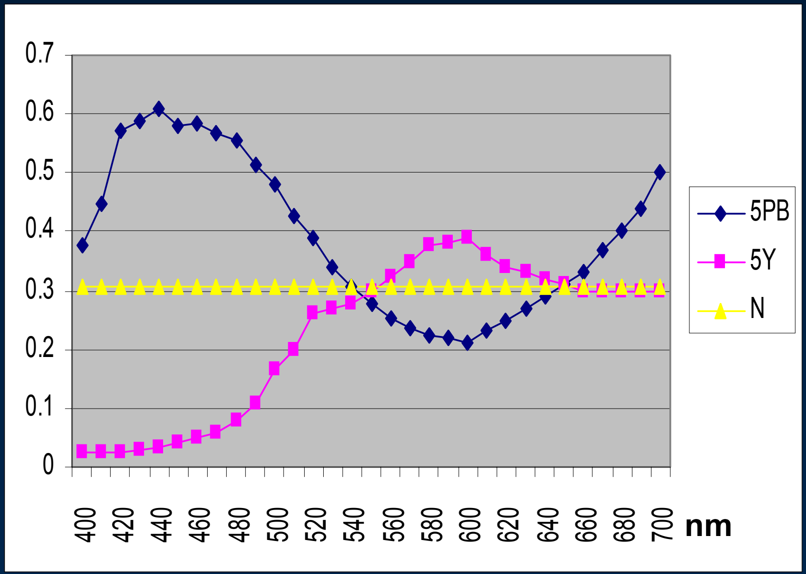
A perceptual color space exemplified by samples can only be reasonably valid for narrowly specified conditions. If the conditions change: observer, surround, lighting, experimental design, etc., the validity declines to a smaller or larger degree.

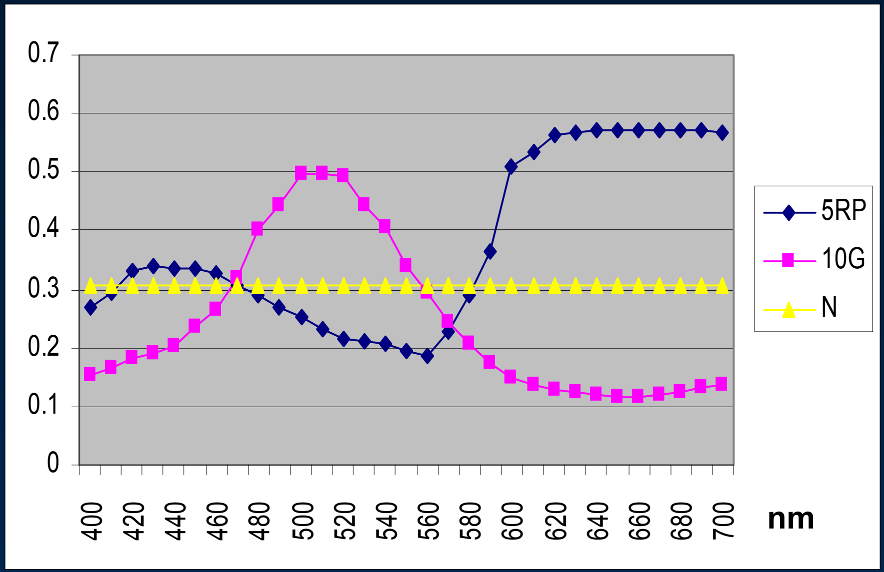
From two to three cones (an interlude)

- Genetic analysis indicates that an "ancient" chromatic system, consisting of *L* and *S* cones, developed approximately 250 million years ago. An *L-S* opponent system resulted in two different hues, presumably blue and yellow. All reflectances of a constant chroma/constant lightness Munsell hue circle were seen either as
  - blue of varying chroma
  - yellow of varying chroma
  - achromatic

- Maximum chroma of the blue color occurred near Munsell hue 5PB, of the yellow color near hue 5Y.

- In the dichromatic system many more reflectances were indistinguishable than in the trichromatic.





Left: Munsell spectral functions with (relatively) largest perceptual distances in the 'ancient system.'

Right: Munsell spectral functions indistinguishable in the 'ancient system.'

Adding a third cone type

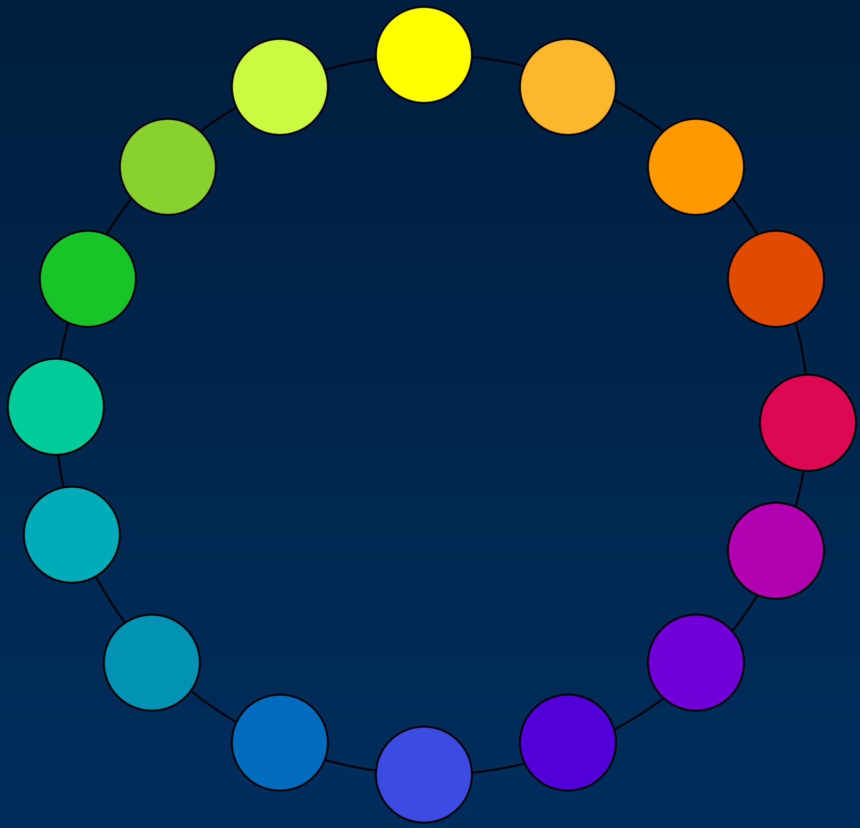
• Addition of a third cone type to the vision system of early primates about 35 million years ago meant:

- addition of a third dimension reduction function, creating a three-dimensional space of spectral power functions
- expanding well distinguishable hues from 2 to 40

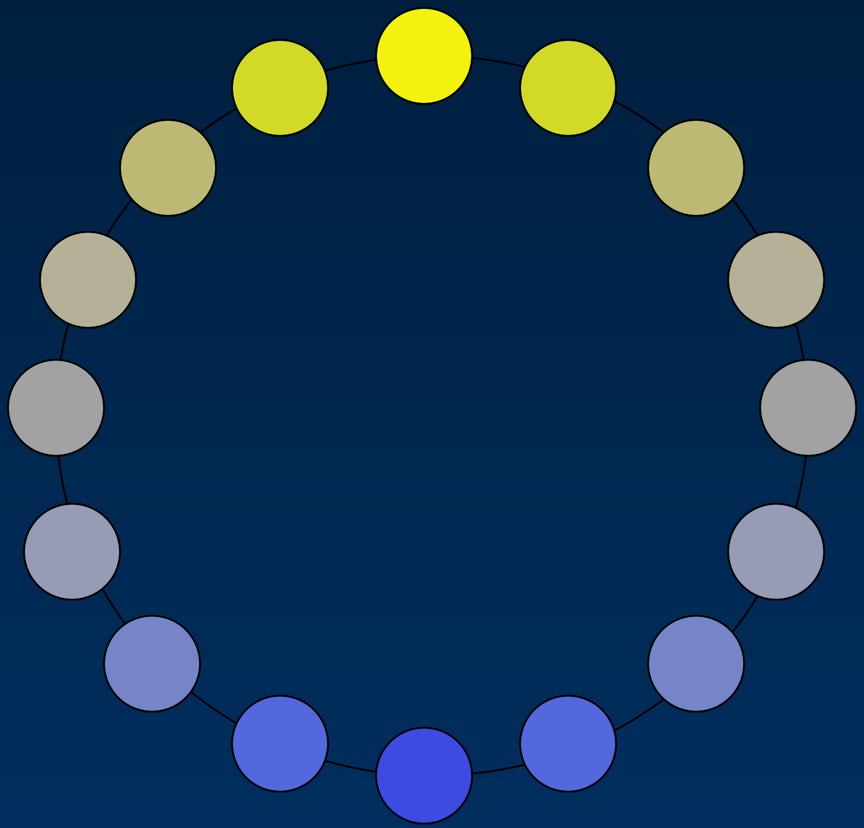
• Assuming 10 chroma steps (of the same perceptual magnitude as the hue steps) and 10 lightness steps, distinguishable color perceptions (at the level of a Munsell 40 hue step) number
  - for the dichromatic ancient system 200
  - for the trichromatic system 4000

a factor of 20, a huge payoff for a moderate investment.

Trichromatic hue circle



Dichromatic version of the same stimuli

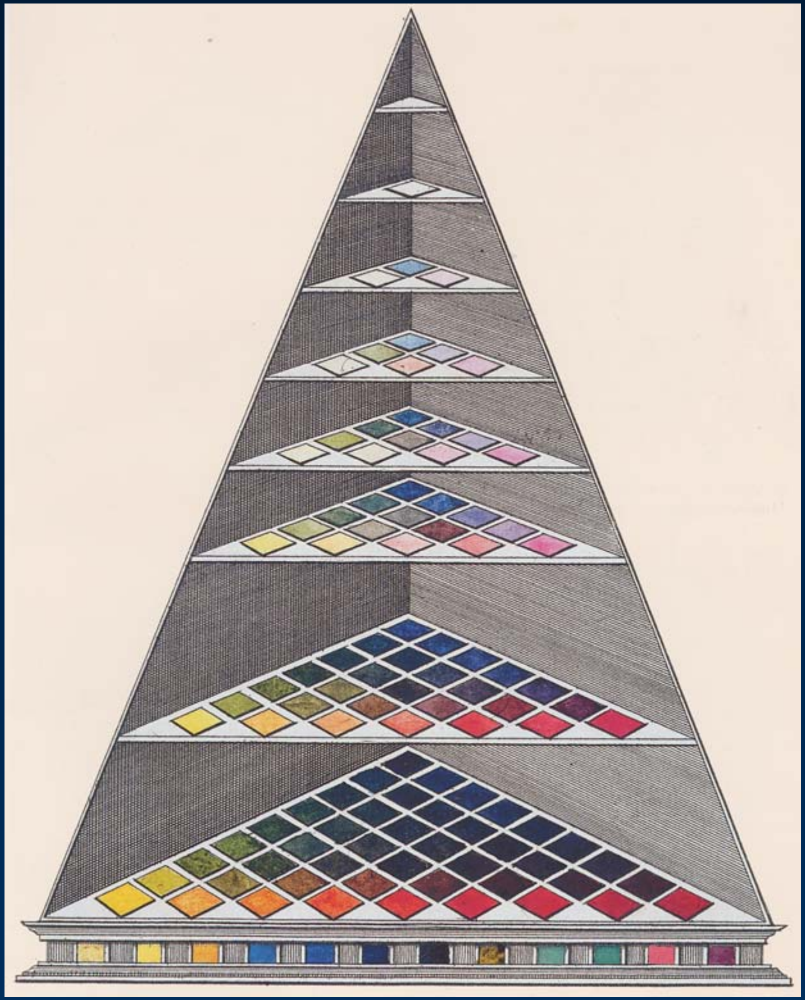


B

Colorant and light source mixture spaces

Colorant mixture spaces

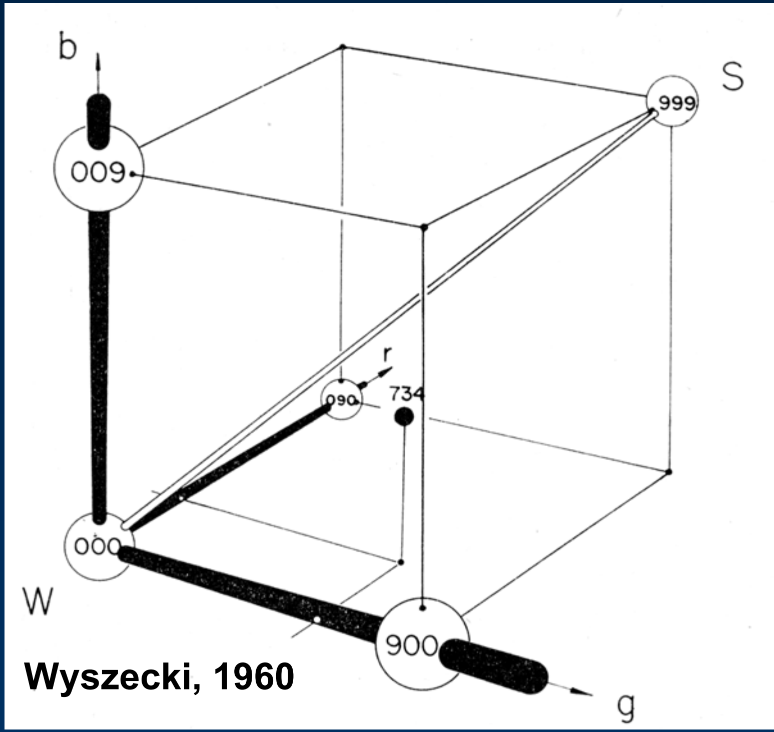
• Historically, a considerable number of colorant mixture spaces have been developed. They generally involve three basic colorants: usually yellow, red and blue. The first of these was Lambert's *Farbenpyramide* of 1772, based on the colorants gamboge, carmine, and Prussian blue.



• Colorant mixture spaces are based on some kind of systematic variation of colorant concentrations. Often a degree of uniformity of perceptual step size was attempted (uniformity in all directions is impossible with systematic variation of concentrations).

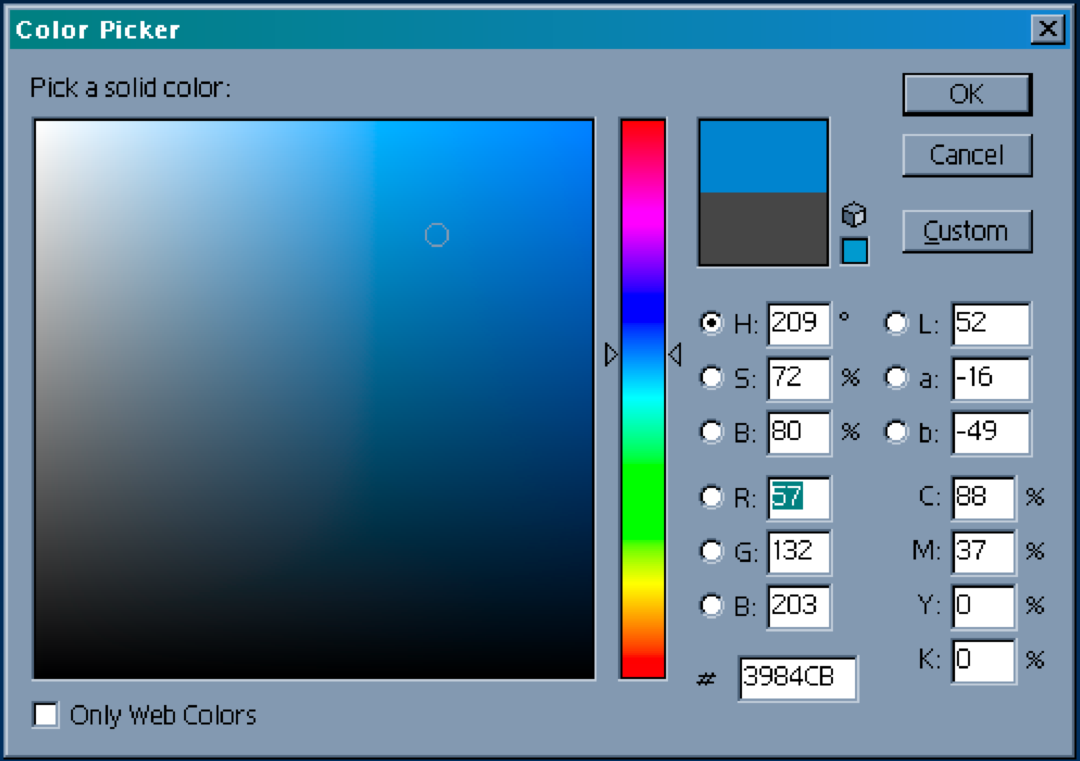
• Some colorant mixture spaces are based on half-tone printing using the printing primaries yellow, magenta, and cyan.

Well known among these is the Hickethier atlas of 1974 with 1000 printed samples arranged in cube form.



Wyszecki, 1960

A kind of cubic mixture space using lights is the modified RGB space of the Adobe Color Picker®. One dimension is a hue scale, another a sort of brightness scale, and the third saturation. The RGB primaries are the phosphors of the video display unit. The display colors are limited to those within the gamut of its primaries.



Nonlinear relation between colorant concentrations and perceived color

The relationship between colorant concentrations and resulting perceived colors (‘standard’ observer and viewing conditions) is in terms of perceptual scales not linear. (The laws of subtractive mixture or a combination of additive and subtractive mixture apply.)

The same applies to electrical input and perceived colors on the video display. (Here the laws of additive mixture apply.)

Additional medium-specific calculations are necessary to generate stimuli in closer agreement with uniform steps of perceived color.

Colorant and emitted light mixture spaces are of considerable technological and practical use.

C

Empirical perceptual spaces

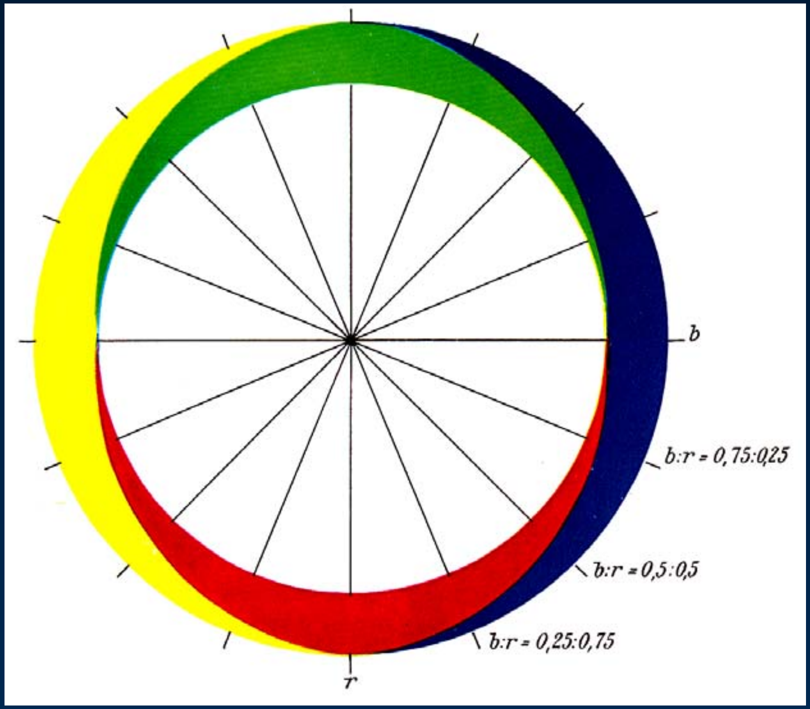
- Hering type
- Attempted isotropic

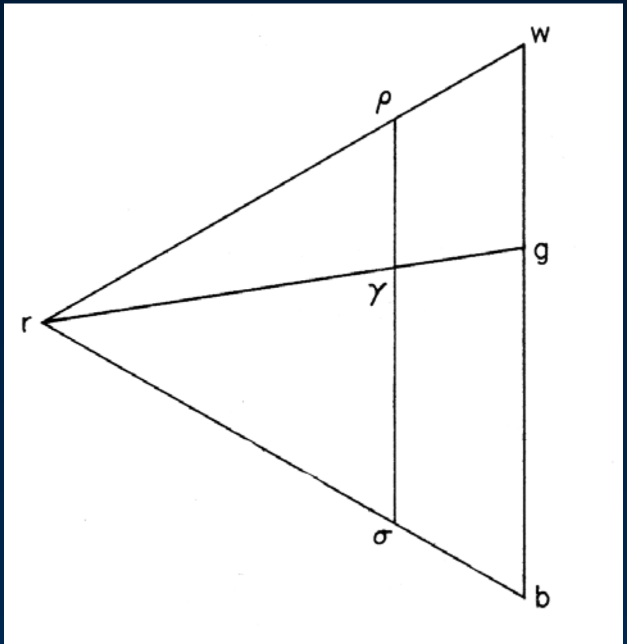
Hering-type perceptual spaces

- In the second half of the 19th century Ewald Hering proposed his opponent color theory and the 'Natural color system.'

- Perceived colors are sums of one or two (of four) chromatic opponent colors, white, and black. The two chromatic opponent pairs are yellow and blue, and red and green.

- Intermediate hues are formed by mixture of two adjacent fundamental colors (experiences). All possible mixtures of a chromatic *Vollfarbe* (full color) with black and white are located on an equilateral triangle.





Left: **Hering's explanation of the formation of the hues of a hue circle from the four chromatic primaries having unique hues.**
Right: Hering's constant hue triangle with full color *r*, white *w* and black *b* at the corners. Arranging the triangles in hue order with a common central gray axis forms a double cone space.

The Swedish Natural Color System (NCS)

A faithful practical implementation of Hering's ideas is the NCS, with an atlas with 1750 color chips.

Image: Swedish Colour Centre

There are 40 constant hue triangles spaced at 10 'percent' change of the four fundamental hues.

Issues with the Hering system

• The vertical dimension has no perceptually uniform meaning: each full color has a different perceived lightness, as Hering well understood.

• Hering assumed that the chromatic power of all full colors is identical, thus placing them on a circle.

• The perceptual distances between unique hues are different. Estimates vary by experiment. Mean estimates in terms of hue angle segments of a circle are: R-Y 75° (65-90), Y-G 83° (75-98), G-B 80° (60-96), B-R 122° (109-135).

As a result the Hering system is far from perceptually uniform (in the sense of isotropic). The designation ‘Natural color system’ seems presumptuous given the arbitrary decisions made.

Perceptually uniform (isotropic) spaces

• Prime examples are the Munsell color solid and the Optical Society of America Uniform Color Scales.

• But as Judd showed in the 1960s a perceptually uniform (isotropic space) in Euclidean form is not possible.
The reason is the phenomenon he termed ‘hue superimportance.’

• Hue superimportance: A hue circle in an isotropic space has a total hue angle of approximately 720° degrees, that is, twice that of a circle. There is a wide variety of experimental evidence. All unit difference contours illustrated in a Euclidean space are elongated, with a diameter ratio of about 2:1.

• There is no common geometrical form for the space curvature implied by this ratio.

• Hue superimportance means that under comparable conditions the change in cone absorptions required to signal a unit difference in saturation is twice that for a unit hue difference and even more for a lightness difference.

• While a mathematical Euclidean model of isotropic color space is not possible calculation techniques have been developed to get around the problem for purposes of color difference calculation.

• Munsell color space is not isotropic but only approximately uniform in terms of its three attributes.

• The only, ambitious attempt at an isotropic space, the OSA-UCS system, is not isotropic because the perceptual results were modified to fit into a Euclidean space. Its experimental data fully confirm hue superimportance.

Connecting perceptual and psychophysical scales

- Attempts to accurately predict perceptual difference from stimulus measurements date back to the early 20th century.

- They were fueled by industrial and technical interests.

- There are several major issues requiring resolution:
  - Psychophysical and perceptual scales are non-linearly related (how?).
  - There are several perceptual effects that add additional complexity to the relationship between perceptual and psychophysical scales.
  - Different scaling methods result in different perceptual scales.
- The results are empirical fittings with no implied neuro-physiological basis.

Non-linearity

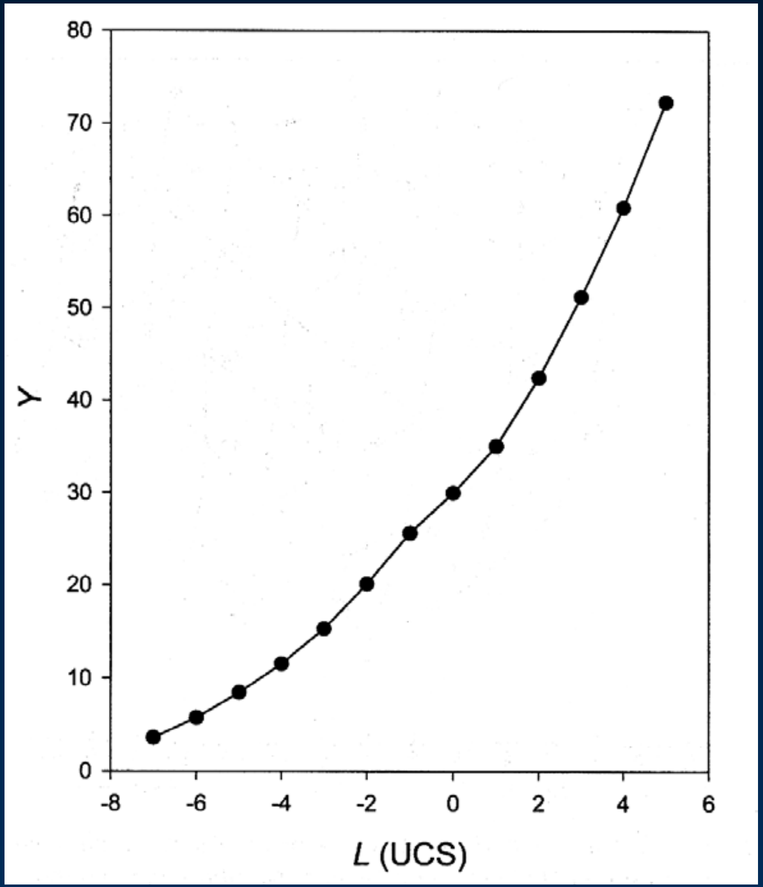
• Differences can be established at the just noticeable difference level or at the unit level of larger differences. The size of the unit difference affects the degree of non-linearity.

• All require a power compression of the stimulus scale: the smaller the difference the closer the power is to 0.
  - The power for lightness scales differs from about 0.10 to about 0.75.
  - Hue scales are non-linearly related with hue angle in a complex way.
  - Power for chroma scales varies by quadrant from about 0.1 to about 0.7.

The Weber-Fechner logarithmic compression has no application in color science.

Perceptual effects

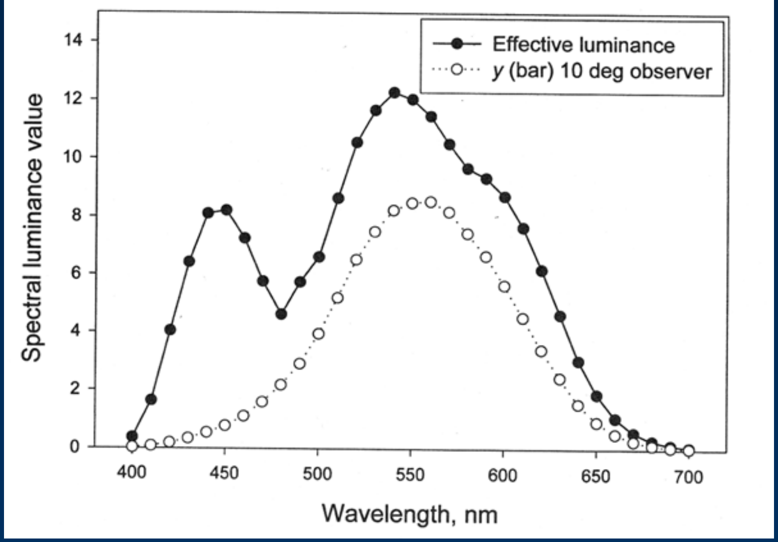
• ***Lightness crispening***:
If the surround lightness falls within the scale (I.e., it is neither white nor black) a Schönfelder (lightness crispening) effect applies: The stimulus change required for a unit perceptual increment or decrement is smallest if the two fields compared straddle the surround in brightness or lightness.
Lightness crispening applies regardless of the size of the lightness unit difference.



Relationship between OSA-UCS perceptual lightness and luminous reflectance. The discontinuity is the result of the lightness crispening effect.

- ***Helmholtz Kohlrausch effect***

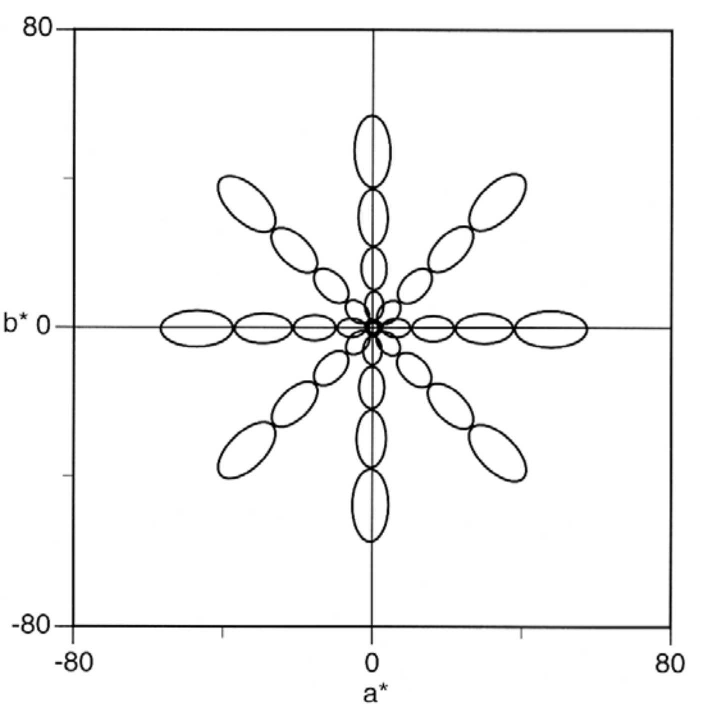
The CIE luminance (*Y*) function does not predict perceived brightness or lightness of chromatic stimuli accurately because of what is known as the Helmholtz-Kohlrausch effect. The perceived lightness of samples not only depends on their luminous reflectance but also on their hue and chromaticity. The magnitude of the effect depends on the experimental conditions (different experiments produce different results). The H-K effect has only been considered in OSA-UCS.



For one particular experiment the full dots illustrate effective luminance, compared
to the CIE standard observer luminance

- *Chromatic crispening*

Crispening also applies to chromatic increments: The smallest chromatic increment is required if the chromaticities of the fields compared straddle that of the surround. For a gray surround the unit contours in a psychophysical diagram are therefore smallest for colors at or near the surround.They grow
In size as the distance increases. Chromatic crispening fades as a function of size of unit difference. At the level of Munsell system differences it has disappeared.



Psychophysical chromatic diagram with conceptul unit difference contours for various hues and a neutral gray background

- *Elongation of unit contour*

Experimental data indicate a (small) unit contour elongation effect as a function of size of difference and of chroma:

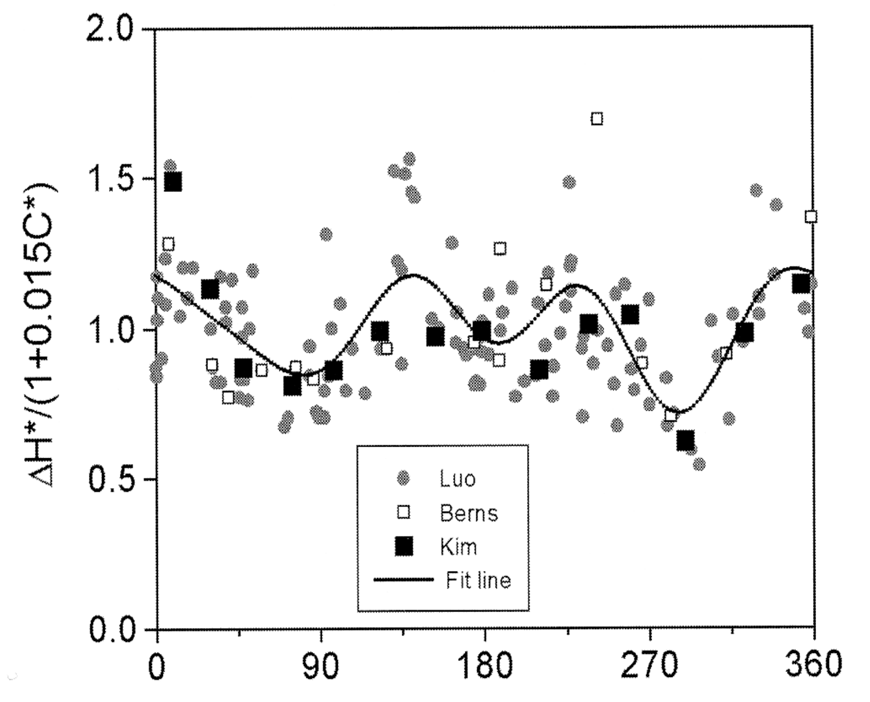
- the smaller the unit difference the more elongated the contour: approximate range 2.7:1 to 2:1

- unit contours close to the surround color are less elongated: their ratio is approximately 1.5:1

- ***Hue angle effect***

Unit hue difference varies as a function of hue angle in the psychophysical color space

Experimental data and fitted function adjusting calculated CIELAB hue differences to be in good agreement with perceptual hue differences.



Cui et al., 2002

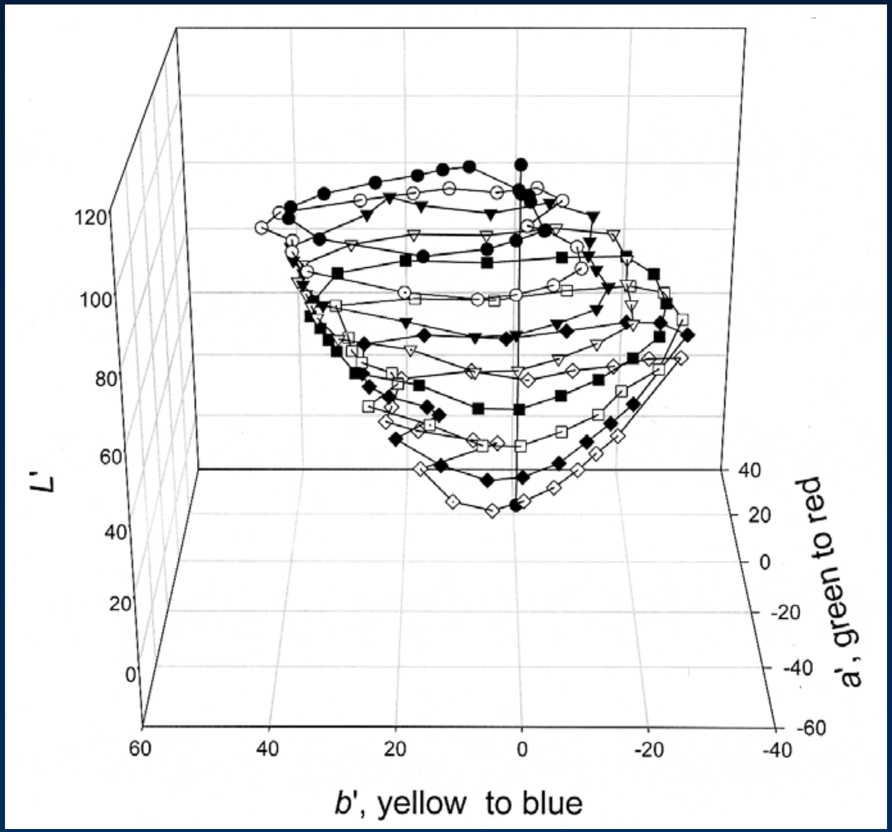
- *Chroma scales*

Depending on the system or experiment *powers* required to linearize psychophysical increments to be in good agreement with perceptual results vary by quadrant:

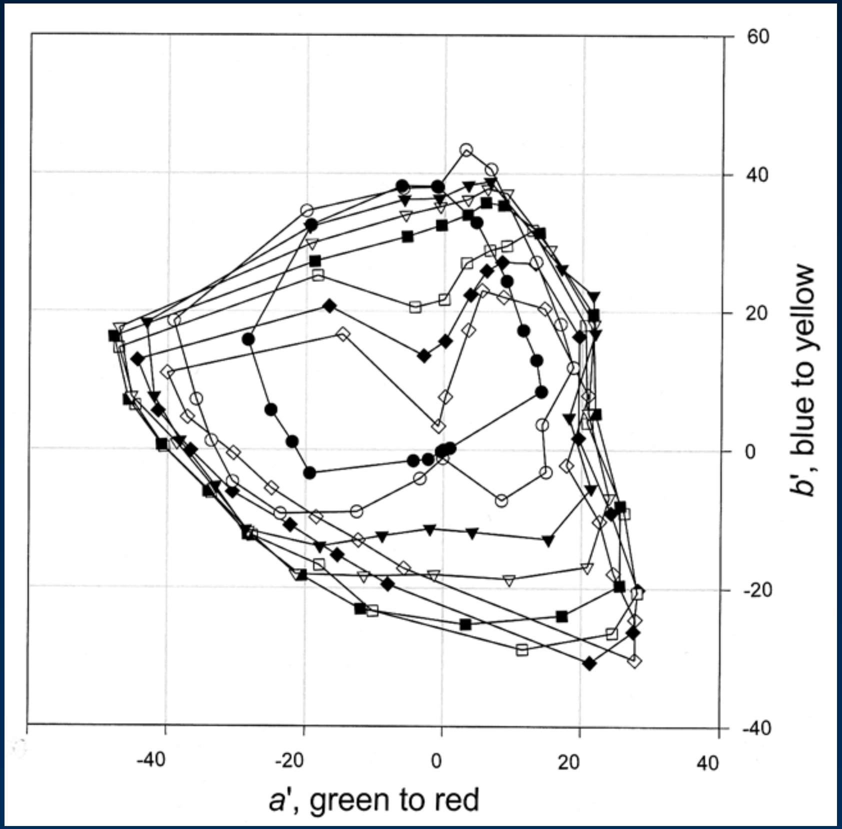
TABLE 5-2 Optimal powers to linearize relations between psychophysical and psychological chroma scale values at or near the axes

Data	*a*+	*a*–	*b*+	*b*–
Munsell Renotations	0.15	0.19	0.42	0.06
COV, %	6.5	4.7	13.1	4.3
Munsell Re-renotations	0.45	0.75	0.70	0.38
COV, %	2.2	6.4	11.1	6.1
OSA-UCS	0.70	0.84	0.58	0.33
COV, %	0.3	0.4	1.8	0.2

Euclidean psychophysical model of isotropic color space at the level of small suprathreshold color differences



View against the *b* axis

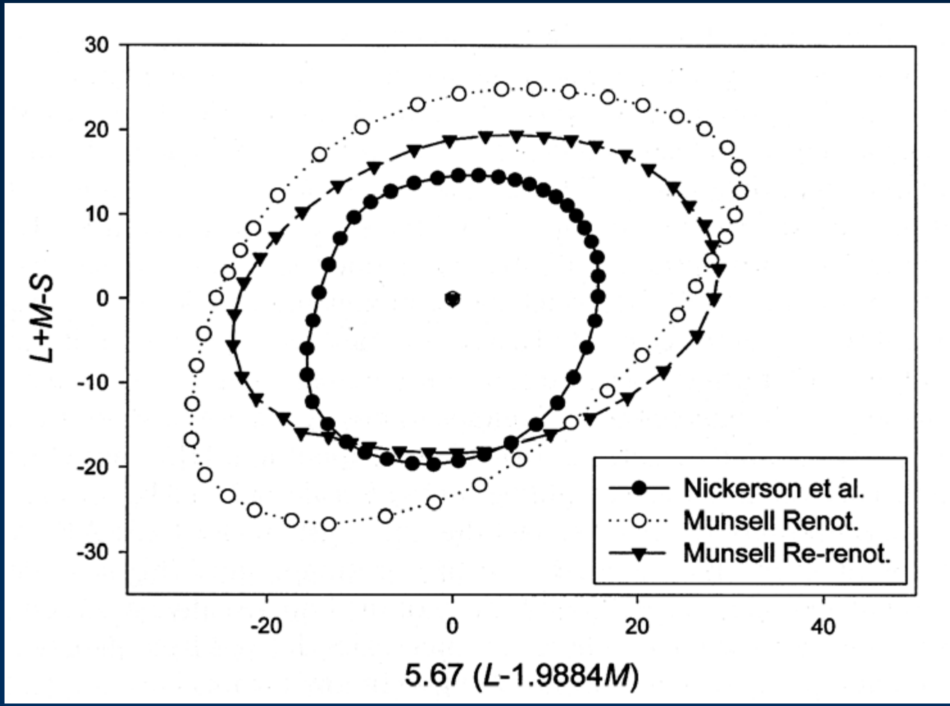


View from the top

Lightness and chromatic crispening and Helmholtz-Kohlrausch effect included; CIE 2° standard observer, equal energy illuminant

Variability in Data

Different large scale experiments in scaling hue and chroma as well as total color differences have resulted in considerable variation in results:



Results of hue scaling around a hue circle and constant chroma scaling in three large-scale experiments (over 10,000 observations each) in a cone sensitivity based chromatic diagram.

Variability in Data, continued

- Accuracy in predicting perceived size of small differences varies significantly for various data sets and by formula.
- For the best formulas the accuracy is approximately 65%.

The causes of variability are *not known*. Among the likely candidates are

- differences in composition of observer panels
- differences in observation conditions: sample size, sample presentation, surround conditions, illumination
- differences in scaling methods.

Replication of experiments is so far lacking.

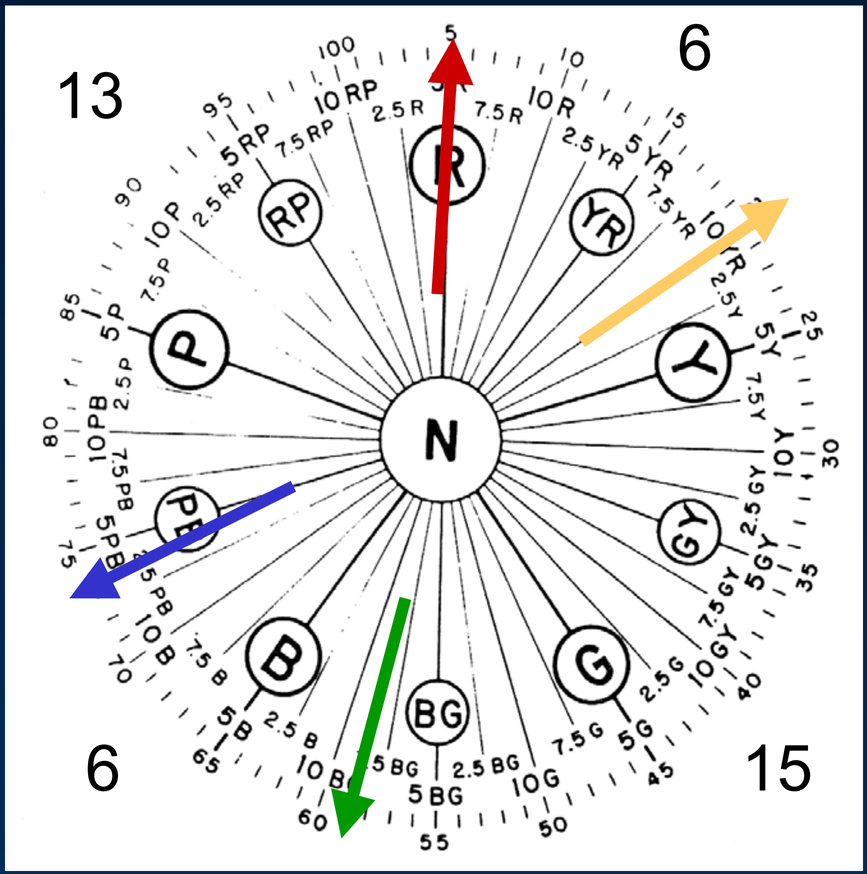
Observer Variability

- Among the possible causes of variability in results of color normal observers are:

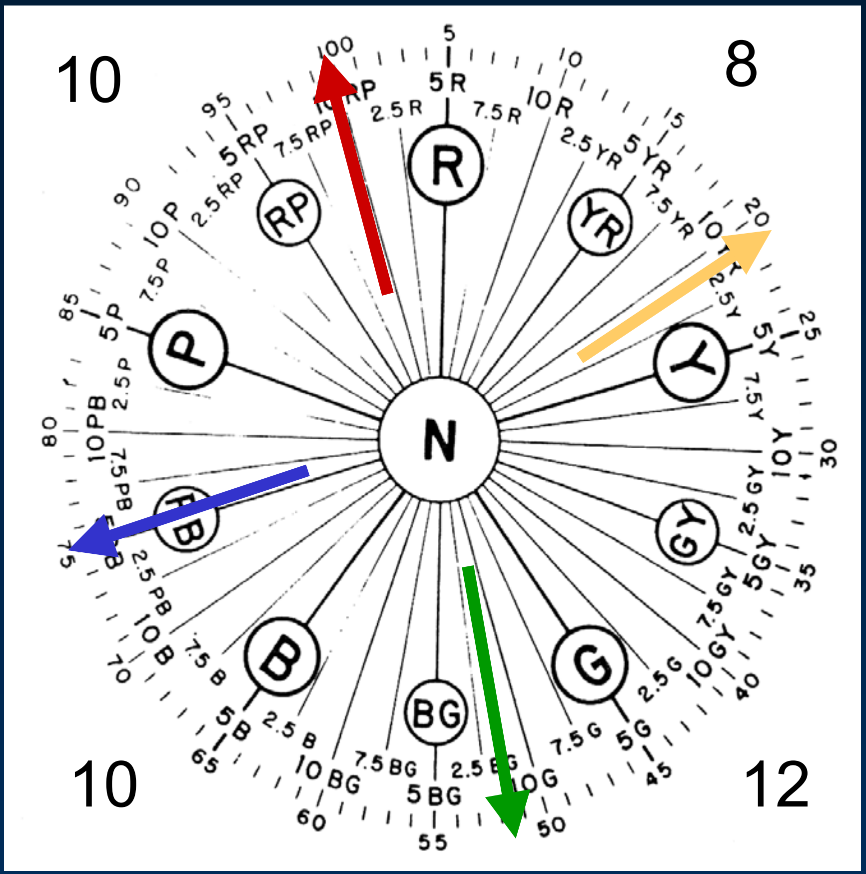
  - differences in color matching functions
  - differences in selection of unique hue stimuli
  - differences in number of cone types
  - subjective conscious or unconscious judgment strategies

The idea of averaging data may be attractive but, given the situation, unlikely to give meaningful results. An averaged color space is *reasonably valid only for 5-10%* of the population.

Unique hue selections for two color normal observers



F, 49



F, 27

Conclusions I

• There are four classes of color spaces:
- stimulus spaces (psychophysical, linear)
- colorant and light source mixture spaces
- perceptual spaces
    - non-isotropic
    - attempted isotropic
- psychophysical spaces optimized to perceptual spaces (non-linear).

• Dimension reduction analysis of color stimuli using three mathematically derived or cone sensitivity functions results in roughly cylindrical spaces in agreement at the ordinal level with perceptual scales, because ...

• This form is implicit in dimension reduction with filters having rounded forms. Two of the optimal three functions have positive and negative values resulting in normalization, as is the case with experimental psychological opponent color functions.

Conclusions II

• There is a considerable level of recovery of information in spectral narrow and broad band stimuli.

• A second opponent system made possible by the third filter (M cone) increased information retention only modestly but discriminability by a factor 20, a large payoff for a modest investment.

• Several kinds of stimulus spaces for a standard observer have been derived, all containing the same information in different reference frames.

• Euclidean isotropic perceptual spaces are not possible.

• The relationship between perceptual and psychophysical spaces is complex non-linear. The fitting process is empirical.

Conclusions III

- Perceptual effects such as lightness and chromatic crispening and the Helmholtz-Kohlrausch effect (themselves variable) add to the complexity.

- Historical perceptual data show large variability. Exact replications have not been performed.

- The causes of variability are unknown but there are several strong candidates.

- Color spaces are only reasonably valid for narrow definitions of observer and experimental conditions. The robustness of such spaces is unknown.

10-30-03