Superior discrimination for hue than for saturation and an explanation in terms of correlated neural noise

M. V. Danilova\textsuperscript{1,2} and J. D. Mollon\textsuperscript{2}

\textsuperscript{1}Laboratory of Visual Physiology, I. P. Pavlov Institute of Physiology, Nab. Makarova 6, St Petersburg 199034, Russia
\textsuperscript{2}Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

The precision of human colour discrimination depends on the region of colour space in which measurements are made and on the direction in which the compared colours—the discriminanda—differ. Working in a MacLeod–Boynton chromaticity diagram scaled so that thresholds at the white point were equal for the two axes, we made measurements at reference points lying on lines that passed at 45° or −45° through the white point. At a given reference chromaticity, we measured thresholds either for saturation (i.e. for discriminanda lying radially along the line passing through the white point) or for hue (i.e. for discriminanda lying on a tangent of a circle passing through the reference point and centred on the white point). The discriminanda always straddled the reference point in chromaticity. The attraction of this arrangement is that the two thresholds can be expressed in common units. All that differs between saturation and hue measurements is the phase with which the short-wave signal is combined with the long-/middle-wave signal. Except for chromaticities very close to the white point, saturation thresholds were systematically higher than hue thresholds. We offer a possible explanation in terms of correlated neural noise.

1. Introduction

In the normal human retina, there are three classes of cone, maximally sensitive in different spectral regions [1]; and our perception of colour depends on neural comparisons of the rates at which photons are absorbed in the different classes [2]. Given three univariant signals, and thus two independent ratios, all visible colours can be represented on a two-dimensional surface. An example of such a ‘chromaticity diagram’ is the MacLeod–Boynton chromaticity diagram (figure 1), whose ordinates are \( L/(L + M) \) and \( S/(L + M) \), where \( L, M \) and \( S \) are the excitations of the long-, middle- and short-wave cones, respectively [3]. The two ordinates of this diagram have physiological counterparts: they correspond to the signals extracted by retinal ganglion cells of the \textit{midget} and the \textit{small bistriatified} types [2].

Although a chromaticity diagram represents all colours, such diagrams do not accurately predict the discriminability of pairs of colours. A pair of chromaticities separated by a given geometrical distance in the diagram may or may not have the same discriminability as a second pair separated by the same distance, depending on the region of the diagram in which the paired chromaticities lie, the direction in which the individual chromaticities differ and the state of adaptation of the eye [3–6]. Yet, in many civil and commercial domains, it is important to be able to predict when two samples will be noticeably different in colour; and several linear and nonlinear transformations of the CIE\textsubscript{1931} chromaticity diagram have been introduced, in successive attempts to achieve a ‘uniform colour space’ in which pairs of colours that are equally discriminable are separated by equal distances [7,8].

(a) The super-importance of hue differences and the problem of Mongean noise

In this study, we compare two fundamental subtypes of colour discrimination: discrimination of saturation and discrimination of hue. We define the two types...
of threshold in terms of the geometry of the chromaticity diagram and not in terms of subjective appearances (so we neglect the Abney effect [9]). We define the white point in the diagram as the chromaticity of the neutral illumination to which the eye is currently adapted (here, metameric to Daylight Illuminant D65). A pair of lights that differ in saturation are lights that lie along a line radiating from the white point. A circular set of chromaticities centred on the white point are described as varying in hue (figure 1a), and the hue thresholds measured here are along tangents to such a circle.

Rather little known in the modern literature on visual science, but much discussed in the mid-twentieth century by those concerned with practical tolerances for surface colours, is a phenomenon that Judd called the ‘super-importance of hue differences’ [7,8,10]. Colour differences measured along a radial line in colour space (i.e. saturation differences) are smaller than the ones that would be expected from differences in an orthogonal direction (i.e. hue differences). If we construct a circle in colour space that is centred on the white point and has a radius of units of perceptual distance, then it is found that the circumference—a hue circle—does not have the same units of perceptual distance but more like 4 π units. Judd concluded that there was no possible Euclidean representation of colour space in which equal distances corresponded to equal discriminability [10].

Judd was primarily concerned with supra-threshold differences, but there is some indication that the super-importance of hue also applies at threshold [11]. It is curious that hue thresholds should be lower than saturation thresholds in this way. The neural channels that extract chromatic information at early stages of the visual system are usually thought of as signalling saturation. The channel is in equilibrium at the adapting chromaticity and it signals departures from this neutral state [12]. Thus, in a conventional view there might be four classes of units in the lateral geniculate nucleus, signalling either increments or decrements along the two cardinal axes of chromaticity space (figure 1b) [13]. To derive precise estimates of hues at angles between the cardinal axes, some form of comparator would be required more centrally to extract the ratios of the saturation signals generated in the distal channels. Estimates of saturation would depend on the vector sum of the same signals, and so it is not obvious why

psychophysical thresholds for hue should be substantially lower than those for saturation. Below, we offer a possible explanation of this paradox.

The classical evidence for the super-importance of hue differences was largely derived from the viewing of surface colours, typically Munsell samples. Here, there is the possibility that physical noise contributes to any difference in discrimination. Real-world surfaces reflect to the eye a specular component (representing the illuminant) as well as a component (the body colour) that derives from selective absorption by pigment molecules within the material. This was first made explicit by Gaspard Monge in 1789 [14,15]; and in modern accounts of colour constancy [16,17], it has been identified as a way in which the visual system could recover the chromaticity of the illuminant by what could be called ‘chromatic triangulation’. Even a matte surface is represented not by a point in chromaticity space but by a distribution of chromaticities that extends from the body colour towards the chromaticity of the illuminant. Moreover, this variation will increase if the eye position is not fixed and changes occur in the angle of viewing the surface. We suggest the term ‘Mongean noise’ for this strictly physical noise. Because Mongean noise lies on a radial line in chromaticity space, it should raise thresholds for saturation but not for hue. It thus could contribute to the ‘super-importance of hue’ when discrimination is measured with material surfaces.

(b) The present measurements

In the work reported here, we use a self-luminous CRT display to compare hue discrimination and saturation discrimination while avoiding the problem of Mongean noise. We measure thresholds for hue and for saturation at reference chromaticities that lie on +45° and −45° radial lines passing through the white point in the MacLeod–Boynton diagram (figure 1a,b). Saturation measurements are made radially through the reference chromaticity, and hue measurements orthogonally through the same point. For each line, in each quadrant of the diagram, we make measurements at different distances from the neutral point.

An essential requirement is a common metric for expressing the two types of threshold. The salient feature of our design is that—for any given reference chromaticity—the
same modulation of the S-cone signal is combined with the same modulation of the \(L/(L+M)\) signal, but in different phases for saturation and for hue. Thus, in the upper right quadrant of the MacLeod–Boynton diagram (Q1), saturation is measured when an increment in \(S/(L+M)\) is paired with an increment in \(L/(L+M)\), and hue is measured by pairing an increment in \(S/(L+M)\) with a decrement in \(L/(L+M)\). The target and distractor stimuli always lay symmetrically on opposite sides of the reference stimulus. The target always incorporated an increment in the \(S/(L+M)\) signal and the distractors incorporated a decrement in \(S/(L+M)\).

The scaling of the vertical axis of a MacLeod–Boynton diagram is arbitrary [18] and the angle taken as 45° will depend on this scaling. We followed recent tradition in the study of chromatic discrimination [19,20] and scaled the \(S/(L+M)\) ordinate so that thresholds for excursions around the white point were equal for the two cardinal axes. Thus, when below we refer to hue thresholds we formally mean thresholds measured along a line orthogonal to a radial line at 45° or −45° in the scaled space (figure 1).

In classical studies of colour discrimination, the observer was often able to inspect the surfaces or lights for an extended period before making a judgement (e.g. [11,21]), and this would allow the chromatic channels of the visual system to adapt towards the current chromaticity—‘self-adaptation’ to the target stimulus. Thus, measurements in different parts of chromaticity space might be made in different states of adaptation. This would be appropriate for setting tolerances for real-world surfaces, where viewing conditions would be similar; but it is inappropriate if the purpose is to analyse the visual mechanisms of discrimination. In this study, we set out to minimally perturb the observer’s adaptation: a neutral adapting field was continuously present, and discrimination was probed with a brief array of target and distractor stimuli.

2. Material and methods
(a) Observers
The observers comprised the authors and three other, female observers who were highly practised but were naive as to the purpose of the experiments. All had normal colour vision as tested by the Cambridge Colour Test [22]. All observers gave informed consent according to the Declaration of Helsinki.

(b) Apparatus and stimuli
Two of the observers were tested in St Petersburg (Russia) and three in Cambridge (UK) using the same experimental programmes. In both laboratories, stimuli were presented on calibrated Mitsubishi CRT monitors (Diamond Pro 2070), controlled by Cambridge Research Systems (CRS) graphics systems (Visage in St Petersburg, VSG 2/3 in Cambridge). For calibration procedures and monitor settings, see [23].

The monitor was viewed binocularly from 57 cm. The steady white background field had a luminance of 10 cd m\(^{-2}\) and its chromaticity was that of Illuminant D65 [6]. A diamond array of black points guided fixation. The stimulus array was presented to the fovea and consisted of a disc divided into four sectors (inset, figure 1b). One sector, chosen at random on each presentation, was the target (T) and the remaining sectors were distractors (D). The array subtended 2° of visual angle.

The duration of the array was 200 ms. This value is a compromise between the self-adaptation that may occur at long exposures (see above) and the ‘tachistoscopic tritanopia’ known to affect very brief presentations [24]. In preliminary measurements, along the two cardinal axes at the white point, we observed tachistoscopic tritanopia below 200 ms, in that the ratio of \(S/(L+M)\) thresholds to \(L/(L+M)\) thresholds increased markedly. The ratio was relatively stable for durations greater than 200 ms.

We specified chromaticities in a MacLeod–Boynton diagram constructed from the cone sensitivities of DeMarco et al. [25]. The diagram represents a plane of equal luminance for the Judd, Observer, where luminance equals the sum of \(L\) and \(M\)-cone signals [26]. We scaled the vertical ordinate so that thresholds for departures from the white point were equal on the two cardinal axes. In these preliminary measurements, we used the same configuration, duration and procedures as for the main experiments, and obtained thresholds for increments and decrements from the white point along the two axes for three of our observers. For each observer, there were five independent estimates of each threshold. On the basis of the average thresholds, we adopted the scaling factor 1.7054 relative to the ordinate of the classical MacLeod–Boynton diagram [27]. As a check, we repeated the measurements after scaling, obtaining a ratio of 1.02 (s.e.m.: 0.04) between thresholds on the two axes.

The test array had an average luminance 30% greater than that of the background when expressed in the \(L+M\) units of our space; but to ensure that observers could not discriminate the target from the distractors by differences in sensation luminance, we jittered independently the \(L+M\) value of each sector by ±1%.

(c) Procedure
Thresholds were measured along four lines radial to the chromaticity of D65 and at angles of either 45° or −45° (figure 1b). Along each line, measurements were centred on a reference chromaticity that differed from D65 in \(L/(L+M)\) value by 1, 2 or 3%. In the orange and green regions, a 3% shift in \(L+M\) chromaticity that differed from D65 in \(L/(L+M)\) was 79.4% correct [23]. The staircase terminated after 15 reversals, the last 10 being averaged to give the threshold. There were six sets of experimental runs, the first set being treated as practice. Thus, any given threshold for a given observer is based on five independent repetitions.

During each threshold measurement, the reference chromaticity was never itself presented, but the discriminanda lay on a line passing through the reference and straddling the reference value. The chromatic separation of the discriminanda was increased or decreased symmetrically around the reference chromaticity according to the observer’s accuracy. The staircase procedure tracked 79.4% correct [23]. The staircase terminated after 15 reversals, the last 10 being averaged to give the threshold. Within the session, hue and saturation thresholds at a given reference were measured in separate blocks of trials, thus giving a total of eight blocks in one session. The orders of blocks and of sessions were randomized.

In any one experimental session, thresholds were measured for reference stimuli at one distance from D65 on all four lines. Within the session, hue and saturation thresholds at a given reference were measured in separate blocks of trials, thus giving a total of eight blocks in one session. The orders of blocks and of sessions were randomized.

In the bar chart of figure 2, we show illustrative data for an observer who was unaware of the purpose of the experiments. The ordinate shows the difference (Δ) between the \(L/(L+M)\) coordinates of the target and distractors (as measurements are made at 45°, the difference in \(S/(L+M)\) is identical; see Material and methods). Each panel corresponds to measurements made at different distances from D65; and within each panel each pair of thresholds corresponds to one quadrant of the MacLeod–Boynton diagram (see figure 1b for the
numbering of quadrants). In all but one case, the threshold for hue is lower than the threshold for saturation measured at the same reference chromaticity. The absolute size of the thresholds increases with distance from the adapting chromaticity (i.e. as one passes from the first to the third panel), and Q1 (the upper right quadrant) shows the weakest difference between hue and saturation.

Figure 3 shows average results for all observers plotted directly in our scaled MacLeod–Boynton diagram after the manner of W. D. Wright’s dashes [28]: the lengths of the line segments represent the separation of targets and distractors at threshold. The dotted line indicates part of the spectrum locus.

Figure 4 shows ratios of saturation thresholds to hue thresholds for the separate quadrants of the MacLeod–Boynton diagram. All quadrants show an increasing ratio with increasing distance from the white point, but systematic differences between quadrants are apparent. The superiority of hue discrimination is most marked in Q4 (upper left)
and least marked in Q1. However, the latter quadrant shows the strongest change in ratio.

4. Control experiment: effects of luminance jitter

In our main experiment, we introduced a +1% luminance jitter to ensure that observers used only chromatic cues. This was a conservative precaution, as chromatic discrimination typically surpasses luminance discrimination (e.g. [34]), and the small gaps between our fields would favour chromatic rather than luminance discrimination [35]. Is it possible, nevertheless, that the luminance jitter elevated saturation thresholds disproportionately? We performed a control experiment in which different levels of jitter (0, 1, 3 and 5%) were introduced in different blocks of trials. The observers and procedures were otherwise as for the main experiment, except that we tested at only one distance from the white point (2%).

Results for this control experiment are shown in figure 5. The four panels, corresponding to the four quadrants, are inset at their appropriate positions within the MacLeod–Boynton diagram. The ordinates of the inset panels show the distance between the discriminanda at threshold. It is clear that the superiority of hue discrimination survives when no jitter is present and that increasing jitter has similar effects on the two types of threshold. A repeated-measures ANOVA shows (after Greenhouse–Geisser correction) significant effects of hue versus saturation ($F_{1,4} = 71.1, p < 0.001$), jitter ($F_{1,64,6.57} = 29.3, p < 0.001$) and quadrant ($F_{2,54,10.17} = 12.3, p < 0.001$). There were no significant interactions.

5. Hue and saturation thresholds very close to the white point

Figure 4 suggests that the ratio between the two types of threshold grows with distance from the white point. For three observers, we have made supplementary measurements very close to the white point, at reference chromaticities that differ from D65 by 2% in their $L/(L + M)$ coordinate. Error bars show $\pm 1$ s.e.m. and are based on the inter-observer variance.

Figure 5. Average thresholds for hue (triangles) and for saturation (circles) as a function of luminance jitter. The four sets of data are symbolically placed within the appropriate quadrants of the MacLeod–Boynton diagram. These measurements were all made at reference chromaticities that differed from D65 by 2% in their $L/(L + M)$ coordinate. Error bars show $\pm 1$ s.e.m. and are based on the inter-observer variance.

6. Discussion

(a) The superiority of hue discrimination: a hypothesis in terms of correlated neural noise

Our main experiment reveals a general superiority for hue discrimination relative to discrimination of saturation. The
Figure 6. Supplementary measurements of hue and saturation thresholds very close to the white point. The abscissa of each panel shows the distance from the white point expressed as the percentage change in the \( L/(L+M) \) coordinate at the reference chromaticity at which the two types of threshold were measured. The four sets of data are symbolically placed within the appropriate quadrants of the MacLeod–Boytton diagram. Error bars show \( \pm 1 \) s.e.m.

Experimental conditions ensured that the two types of measurement differed only in the signs with which modulations on the cardinal axes were combined, and the modulations were always centred on the same reference chromaticity, so the magnitudes of the two types of thresholds can be directly compared. Clearly, it would not be possible, from threshold measurements made under the present conditions, to construct a Euclidean colour space in which equal geometric distances corresponded to equal discriminability (see Introduction).

Why should thresholds for saturation systematically exceed those for hue? We offer a provisional theory in terms of correlated neural noise within the visual system.

In both the retina and the visual cortex, neurons are known to exhibit correlated variations in excitability [36–38]. Typically, the correlations are attenuated with increasing spatial separation of a given pair of neurons. Of special interest, for our present study, is the demonstration by Greschner and colleagues of correlation in the primate retina between neural firing in small bistratified ganglion cells (which draw ON input from S cones) and that in nearby ON midget cells (which draw ON input from either L or M cones) [38].

Figure 7 illustrates our hypothesis of how correlated neural noise within the visual system. This view was classically expressed by Kuehni, who wrote: 'In practical terms, there appear to be two independent systems: one that assesses changes in the ratio of two opponent color signals (assuming a two-process hue detection system) and the other changes in the size of the vector sum of the color signals (assuming a two-process saturation detection system). The 'super-importance of hue' has sometimes led to the conclusion that hue and saturation are analysed by different systems. This view was classically expressed by Kuehni, who wrote: 'In practical terms, there appear to be two independent systems: one that assesses changes in the ratio of two opponent color signals (assuming a two-process hue detection system) and the other changes in the size of the vector sum of the color signals (assuming a two-process saturation detection system).
The two seemingly operate independently of each other and are not connected in a euclidean sense ([8, p. 331]). Our hypothesis suggests how the two types of discrimination might depend on the same distal signals, although necessarily in one case it is the ratio of those signals that is identified and in the other, it is their sum.

An interesting analogy can be made between hue and saturation, on the one hand, and the auditory attributes of pitch and loudness on the other. In the classification of S. S. Stevens, hue is a *metathetic* sensory dimension like pitch, and saturation is *prothetic* like loudness [39]. In psychoacoustics, it has often been held that frequency discrimination could not depend on the same neural signals as intensity discrimination. The latter is thought to depend on the spike rate of cortical neurons, but the relatively high resolution observed for frequency suggested that pitch discrimination could not depend on the same signal and instead must depend on the timing of spikes. However, Micheyl et al. [40] explain the disparity of the two types of threshold by correlated noise in the same types of neuron.

In the case of vision, a hypothesis analogous to the present one was explored by Regan & Beverley [41] for the case of orientation discrimination. They point out the general principle that ‘Opponent processing has the feature that noise in the first-stage outputs...can be partially or even completely suppressed by the opponent element’ ([41, p. 153]).

(b) Intermediate colour channels?
For the purposes of the model outlined above, we took the primary signals to be $S/(L + M)$ and $L/(L + M)$, but chromatic

---

**Figure 7.** How correlated neural noise could explain the superiority of hue discrimination. In each panel, the $x$- and $y$-axes represent the signals in the two ‘cardinal’ chromatic channels of the early visual system. The left-hand panels correspond to hue discrimination and those on the right correspond to saturation discrimination. (a) No noise is assumed: the responses generated by the target and the distractors plot as points in the two-dimensional space. (b) The case where independent Gaussian noise is present in the two cardinal channels. The circles represent one contour of the probability density function produced by a given chromaticity. (c) It is assumed that noise in the two channels is correlated. Now the probability density distributions are elliptical with positive slopes. When the stimuli differ in saturation (right-hand panel) the distributions for the discriminanda overlap, but when the stimuli differ in hue—when a change on one axis is accompanied by an opposite change on the other axis—the overlap of the probability density distributions is less.
channels tuned to intermediate, non-cardinal directions of colour space have often been postulated (e.g. [42]), channels that are usually taken to be central to the ‘cardinal’ channels but could conceivably be in parallel with them [43]. In principle, an explanation of the type offered above could still apply if hue discrimination depended on a comparison of the signals in channels tuned to intermediate directions in colour space.

Discrimination ellipses oriented obliquely have actually been taken as independent evidence for ‘intermediate’ chromatic channels. Our hypothesis of correlated noise offers an alternative explanation for such ellipses.

(c) Variation between quadrants of the MacLeod–Boynton diagram

For reference chromaticities at moderate distances from the white point, it appears to be a general rule that saturation thresholds exceed those for hue; and above we offered a possible explanation for this result. However, a complete explanation will need to take into account the differences of the strength of the effect in different quadrants of the MacLeod–Boynton diagram (figure 4). For Q2 (lower right) and Q4 (upper left), the ratios are largest, and the effect is maintained in regions very close to the white point (figure 6), whereas the effect is attenuated or reversed near the white point for quadrants Q1 and Q3. These results are consistent with earlier studies [19,20], where discrimination ellipses were measured in the different quadrants of DKL [13] space. A similar pattern is seen in the variance of thresholds for different observers in the results of Elliott et al. [42].

In this context, we note a provocative coincidence. In the present experiment, in equating cardinal axes to give equal thresholds at the white point, we found that we needed to scale the \( S/(L + M) \) ordinate of the MacLeod–Boynton diagram by a factor of approximately 1.71. Independently, in earlier experiments [43], we scaled this ordinate so that a line running between the wavelengths of unique yellow and unique blue ran at \(-45^\circ\); a very similar factor (1.64) was needed. The locus of unique yellows and blues in turn closely coincides with the ‘caerulean line’ in the outside world, the locus comprising mixtures of sunlight and sunglight [15,44].

This means that in quadrants Q2 and Q4, the present saturation thresholds lie approximately along the locus of unique yellows and blues, while hue judgements in these quadrants coincide with the category boundary between reddish and greenish hues. We have earlier noted that an optimum of discrimination occurs in the latter case [43,45] (i.e. where thresholds are being measured orthogonally to the unique-blue locus). It remains possible that this region of enhanced discrimination does correspond to the equilibrium point of an intermediate neural channel that draws opposed inputs from \( M \) cones and from \( L + S \) cones (for discussion, see [43]).

7. Conclusion

Within the area of colour space that we have explored, thresholds for discriminating saturation are systematically higher than those for discriminating hue, except for chromaticities very close to the white point. We offer a possible explanation in terms of correlated neural noise in distal chromatic channels in the visual system. There is empirical evidence that noise in these two channels is indeed correlated [38].

References
