Chromatic detection and discrimination analyzed by a Bayesian classifier

Rhea T. Eskew, Jr. *, Jessica R. Newton, Franco Giulianini

Department of Psychology, Northeastern University, 125-NI, Boston, MA 02115, USA

Abstract

Detection and threshold-level discrimination of Gabor patches were studied under the conditions of noise masking, in an attempt to isolate 'higher-order' or nonclassical color mechanisms. Detection contours in the equiluminant plane of cone contrast space were measured by varying test chromaticity in the presence of chromatic masking noise. Three equiluminant noise directions were used, in separate experiments. In the discrimination experiment, observers had to discriminate between pairs of stimuli that were fixed at their masked threshold contrasts. A Bayesian color classifier model was used to analyze the discrimination data, with no free parameters. There was no evidence of nonclassical color mechanisms in either the detection or discrimination data. © 2001 Elsevier Science Ltd. All rights reserved.

1. Introduction

One of the central contemporary issues in color vision is the number of psychophysical color mechanisms that exist. Much experimental data is consistent with the view that there are but four chromatic mechanisms: red (R), green (G), yellow (Y), and blue (B) , each formed by opponent combinations of cone signals. However, recent evidence from detection (e.g. Krauskopf, Williams, & Heeley, 1982; Krauskopf, Williams, Mandler, & Brown, 1986; D'Zmura, Lennie, & Krauskopf, 1987; Gegenfurtner & Kiper, 1992), visual search (e.g. D'Zmura, 1991) and color appearance (e.g. Webster & Mollon, 1994) tasks has been taken to indicate the existence of additional mechanisms, termed here 'nonclassical mechanisms', tuned to intermediate hues, such as 'orange' or 'blue–green'. These findings are consistent with reports of cortical color cells that are also tuned to noncardinal color directions (Lennie, Krauskopf, & Sclar, 1990; Gegenfurtner, Kiper, & Levitt, 1997; Kiper, Fenstemaker, & Gegenfurtner, 1997). However, in our recent series of noise masking experiments in the (L,M) plane (Giulianini & Eskew, 1998),

we found no evidence of nonclassical mechanisms (see also Sankeralli & Mullen, 1996, 1997).

Li and Lennie (1997), using static noise masks, directly compared the equiluminant and (L,M) planes of color space in a texture segmentation task, and found evidence favoring nonclassical color mechanisms in the former but not the latter plane. This difference between color planes probably has to do with the S cones rather than with equiluminance per se. The anatomy and physiology of the S-cone ON pathway are very different from that of the L and M cone ON and OFF pathways (or S-cone OFF pathway) (Calkins, 1999; Dacey & Lee, 1994), and it might not be surprising to find that S cones contribute to additional mechanisms. Using a transient adaptation procedure, McLellan and Eskew (2000) recently found psychophysical evidence for two distinct S-cone detection mechanisms, presumably ON and OFF pathways, that had different relative amounts of L and M cone opponent input – as if there were nonlinear Y and B mechanisms with asymmetric cone weights.

The present study extends our previous noise masking experiments (Giulianini & Eskew, 1998) to include S-cone modulations (in the equiluminant plane). In addition to the detection procedure, we also used a

---

* Corresponding author.

discrimination procedure and a Bayesian classifier model to determine how many mechanisms were active at threshold in the equiluminant plane. The logic of the discrimination procedure depends upon two closely related assumptions, both of which derive from Müller's Law of Specific Nerve Energies (Boring, 1942): (a) that two stimuli that are detected via a single chromatic mechanism can be made indiscriminable by adjusting their relative strengths (univariance); (b) that two stimuli that are detected via two different chromatic mechanisms are as discriminable as they are detectable. These two assumptions can be summarized by saying that chromatic detection mechanisms are 'labeled lines' (Graham, 1989; Watson & Robson, 1981). In addition, to satisfy a rigorous definition a chromatic mechanism should either have a fixed spectral sensitivity or at least there should be some explicit rule by which the spectral sensitivity is altered (as by adaptation, for example). According to this definition, R and G (and Y and B) are separate cone-opponent mechanisms, not two halves of a single mechanism. Whether R and G (Y and B) are symmetric (equal but opposite cone weights) is a separate question (see below).

In both the detection and discrimination experiments, we used noise masking to desensitize R and G, which are by far the most sensitive mechanisms (Chaparro, Stromeyer, Huang, Kronauer, & Eskew, 1993; Eskew, McLellan, & Giulianini, 1999). We hoped thereby to reveal less-sensitive nonclassical mechanisms tuned to other directions. However, neither the detection nor the discrimination procedure found any evidence for additional mechanisms besides the classical ones.

2. Methods

2.1. Apparatus

A Nanao T560i monitor, driven by a standard video card with 8-bit digital-to-analog converters (DACs), was used to display the stimuli. The mean monitor field was circular (9.4° diameter), white ($X = 0.309$, 0.331), and produced a retinal illuminance of 130 Td. Two background fields (474.4 and 575.3 nm) were optically superposed over the monitor image and combined in a radiance ratio to make them metameric with the monitor white for a Vos-modified Judd observer (Vos, 1978). All fields were viewed through a 2.4 mm diameter optical pupil; total retinal illuminance was 190 or 250 Td in different experiments, with 130 Td provided by the monitor in both cases. Head position was stabilized with a dental impression mouth bite, and a five-element achromatizing lens (Powell, 1981) was used to correct for both lateral and longitudinal chromatic aberrations. Spectroradiometric calibration of the monitor was performed at 1.05 nm intervals over the spectrum. Radio-

metric calibrations of the background fields were performed every 2 h during experimental runs. The monitor was linearized via gamma correction lookup tables.

2.2. Stimuli

Test stimuli were horizontal sine-phase Gabor patches, of 1 cpd and with $\sigma = 1°$. Due to the contrast resolution of the 8-bit DACs a maximum of $\pm 2.1\sigma$ was drawn on the screen. Gabor patches, rather than spots or other unipolar stimuli, were used because Gegenfurtner and Kiper (1992) found evidence for higher-order color mechanisms using Gabors but not using large square unipolar patches. Our stimuli were flashed for 200 ms. The noise was binary; it consisted of flickering horizontal lines, that randomly and independently changed from one chromaticity in the equiluminant plane to a symmetrically-opposite chromaticity (on the other side of the white point). Each line switched chromaticity with probability 1/2 at 16.8 Hz. The noise was presented continuously throughout the experiment run. Test and noise were spatially separated: horizontal test lines of 2.64 min height were interdigitated with horizontal noise lines of the same heights (Giulianini & Eskew, 1998, Fig. 2d). The noise lines were set to the white adapting field when no noise was used. The high spatial frequencies created by this 'halftoning' procedure were invisible. The halftoning effectively halves the size of the smallest contrast step that the device can produce, and the adapting background produces a further reduction; both measures provide finer control of stimulus contrast than would be obtainable using 8-bit DACs and no background. Test and noise strength are measured as the length of the vector of cone contrasts produced by the stimulus at its peak (Eskew et al., 1999; Stromeyer, Cole, & Kronauer, 1985); however, to facilitate comparisons with other studies in which halftoning was not used, we halved all of the peak contrast values (as in Giulianini & Eskew, 1998) and adjusted mechanism sensitivity estimates accordingly.

2.3. Procedure

Detection thresholds were measured with a 2AFC staircase method. After adapting to the steady background (including noise, if used) for 1–2 min, the observer initiated a run of 100 two-temporal interval trials. Two independent staircases were intermixed within a run, which consisted of presentations of a single test (and noise) color direction. Each trial consisted of two 200 ms intervals, signaled by tones, separated by 400 ms. Test contrast was decreased by 0.1 log units after three consecutive correct responses and increased by the same amount after a single error. All of

the frequency of seeing data from a run was accumulated and analyzed by fitting a Weibull function to extract a threshold estimate (corresponding to 82% detection) and a psychometric slope estimate (Watson, 1979). Data figures show mean and standard error, based upon between-run variability. In most cases runs occurred in different sessions on different days. Generally, three runs were used for a given test condition; observer JH's detection thresholds were variable and somewhat irregular, and as many as six runs were used for a few of her test angles (not counting replications, plotted separately).

For the discrimination experiments, two stimuli were fixed at their detection thresholds, and presented in the two intervals of a trial in random order. One color direction was designated as correct (the 'standard'), and the observer learned from the feedback tones how to select the standard stimulus in practice runs. Stimuli potentially could be discriminated either on the basis of chromaticity or phase (i.e. two Gabor patches separated by 180° color angle in the equiluminant plane have the same chromaticities but their carrier sine waves differ by 180° spatial phase — e.g. the red and green bars are interchanged in position).

Six well-practiced observers were used, all of whom had normal color vision; two of these observers also participated in the experiments of Giulianini and Eskew (1998). In the discrimination experiment, observers were given a number of practice runs, especially in conditions in which the discrimination was difficult, in order to insure that performance would not improve with practice. Final discrimination performance was generally estimated from two to four runs.

2.4. The equiluminant plane

For four of the six observers (FG, PK, JRN, and RTE), the $(\Delta M/M)/(\Delta L/L)$ ratio at equiluminance was determined individually; this ratio determines the angle

θ of the equiluminant direction in the (L, M) plane. FG's equiluminant ratio was determined from detection data: we selected the direction in the (L, M) plane of cone contrast space that was parallel to a set of achromatically-detected thresholds obtained in the presence of M-cone masking noise (data shown in Giulianini & Eskew, 1998, Fig. 7a). This direction was close to 120°. For PK, JRN, and RTE, a minimally-distinct border procedure was used (Boynton, 1978). An 'isolated edge' (McLellan, Goodman, & Eskew, 1994), consisting of a horizontal step function multiplied by a circular Gaussian, was flashed for 200 ms every 3 s. The two halves of the stimulus were symmetric about the origin in the (L, M) plane, and the observer could use buttons to adjust the angle of the stimulus in that plane. The observers were asked to minimize the distinctness of the centrally-viewed edge. The contrast of the edge was set to three different levels (the maximum available, and 2/3 and 1/3 of that) in separate runs. Data were analyzed by fitting a line through the origin and through the set of minimally-distinct settings. The $(\Delta M/M)/(\Delta L/L)$ ratio of that line was taken as the equiluminant direction. The ratios were -1, -1.22, and -1.75 for PK, JRN, and RTE, respectively. All of these are less than the Judd V_λ function predicts, but the angular difference is small: the Judd observer's equiluminant direction is at $\theta = 118.1°$ in the (L, M) plane (on our white field), while PK, JRN, and RTE were at $\theta = 135°$, $129.2°$, and $119.5°$, respectively. For the two remaining observers (JDP and JH), $\theta = 120°$ was used, just as for FG.

The equiluminant plane for each observer is represented by a pair of orthonormal basis vectors: the S cone $\{\Delta L/L, \Delta M/M, \Delta S/S\}$ $\{0,0,1\}$ vector, and the unit vector in the observer's θ direction, symbolized as $\{L_e, M_e, 0\}$. For example, for RTE $L_e = 1/\sqrt{4.06}$, and $M_e = -1.75/\sqrt{4.06}$. Angles are calculated for each observer with respect to this basis set. 'Red' is at 0° in the equiluminant plane, and positive S-cone modulations are represented at 90°.

2.5. Noise

Noise directions and strengths are given in Table 1. For the 0°\180° and 90°\270° directions, the contrast was either the maximum possible, or the highest contrast that would produce measurable masked thresholds in all color directions (judged on the basis of pilot experiments). For JDP (Fig. 3) who was quite insensitive along the S-cone direction, we overestimated the useable noise contrast and were not able to measure an S-cone threshold. For the intermediate noise angles used for FG and PK, noise contrasts were adjusted to produce approximately the same projection onto 1° and 181° (last column of Table 1), where we expected to find the vectors of the R and G mechanisms (see below).

Table 1
Noise angles and noise contrasts

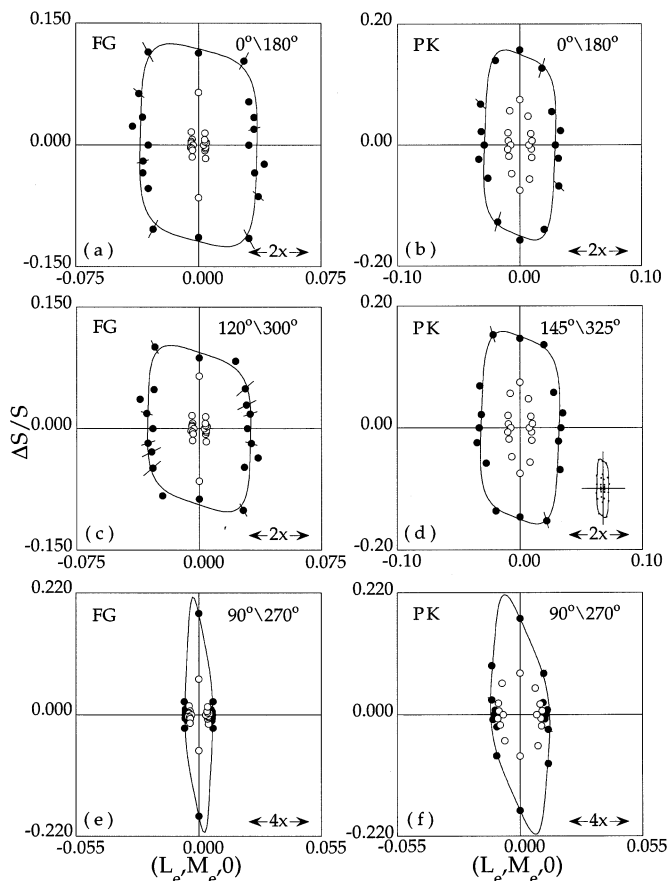| Observer	Noise angle	Noise contrast	Projection on 1°
FG	0°\180°	0.04	0.04
	120°\300°	0.07	0.03
	90°\270°	0.20	0.00
PK	0°\180°	0.02	0.02
	145°\325°	0.03	0.02
	90°\270°	0.16	0.00
JRN	0°\180°	0.05	0.05
JH	0°\180°	0.04	0.04
JDP	0°\180°	0.05	0.05
RTE (Fig. 5)	0°\180°	0.03	0.03
RTE (Fig. 7)	0°\180°	0.04	0.04

Fig. 1. Detection contours in the equiluminant plane of cone contrast space, for two observers (250 Td). The horizontal axis is along the equiluminant $L\backslash M$ direction (L cones positive to the right), and the vertical axis is the S-cone direction. Open symbols are unmasked thresholds, and filled symbols are the masked thresholds, with binary masking noise presented along the $0°\backslash180°$ ($L\backslash M$) direction in panels (a) and (b), the S-cone direction in panels (e) and (f), and an intermediate direction in panels (c) and (d) (see Table 1). In the upper four panels the horizontal axis scale is expanded by a factor of two relative to the vertical axis; it is expanded by a factor of four in the lowest pair of panels. The inset in panel (d) shows the correct aspect ratio of the detection contour, for comparison. Error bars are ± 1 standard error based upon between-run variability. The solid lines are the standard model predictions (standard R and G weights), as described in the text and Table 2.

3. Results and discussion

3.1. Detection results

The open symbols in Fig. 1 show, for two observers, unmasked detection thresholds in the equiluminant plane. The points have been plotted twice, representing each peak of the symmetric Gabor pattern. The horizontal axis has been expanded relative to the vertical axis, as indicated at the lower right corner of the panel. The inset in Fig. 1(d) replots the data with an aspect ratio of 1 to illustrate the true shape of the detection contours: as shown before (e.g., Cole, Hine, & McIlhagga, 1993; Sankeralli & Mullen, 1996), unmasked

contrast thresholds are much higher along the S-cone axis than along the $L\backslash M$ axis, for briefly flashed foveal stimuli.

The filled symbols in Fig. 1 show the masked detection thresholds. In panels (a) and (b), the noise was in the equiluminant $L\backslash M$ ($0°\backslash180°$) direction. The noise raises the threshold for all of the color directions tested. In particular, the long, straight-line flanks, are shifted outward (by $\sim 8X$ for FG and $\sim 3X$ for PK), with no change in slope. This is consistent with detection by a pair of cone-opponent mechanisms (Section 3.3), presumably R and G. Thresholds along the S-cone axis are also elevated, by a factor of 1.8 for FG and 2.1 for PK. Thus, whatever mechanisms are most sensitive to S-cone modulation are not orthogonal to the observer's $L\backslash M$ axis.

Fig. 1(c) and (d) show the same, no-noise thresholds, along with new masked thresholds. For FG, the noise was at $120°\backslash300°$ in this plane; for PK, the noise was at $145°\backslash325°$. In each case, the pattern of results is similar to the masked pattern in panels (a) and (b): the long flanks are pushed out with no change in slope, and there is some masking in the S-cone directions, at least for PK (1.3-fold for FG, and two-fold for PK). The noises at $120°\backslash300°$ and $145°\backslash325°$ produce about the same masking of the R and G flanks of the detection contour as the $0°\backslash180°$ noise does, as they should if the R and G mechanisms are near 1° and 181°, respectively (Table 1); this result is unlikely if there are additional mechanisms in this region of the plane.

We also examined the effect of S-cone noise, in part to shed light on a controversial issue: whether S cones contribute to the R and G *detection* mechanisms, as they do for the red and green hue mechanisms (Boynton, Nagy, & Olson, 1983; Mollon & Cavonius, 1987; Stromeyer & Lee, 1988). Fig. 1(e) and (f) show that pure S-cone noise substantially masks the S-cone test, and slightly masks the reddish/greenish tests near the horizontal axis. This second result is consistent with a small S-cone input to R and G (Eskew et al., 1999; Krauskopf et al., 1982; Sankeralli & Mullen, 1996; Stromeyer, Chaparro, Rodriguez, Chen, Hu, & Kronauer, 1998). The masking of R and G mechanisms by S cone noise is much less than the masking of Y and B mechanisms by $L\backslash M$ noise, indicating that the contribution of the long-wave cones to Y and B is substantially more than the contribution of S-ones to R and G.

Threshold elevations along the vertical, S-cone direction in the panels of Fig. 1 vary from a minimum of 1.3-fold to a maximum of 2.8-fold (both for FG). One might have expected a larger variation, given that the noise angle changed by 90° and given the high contrast noise used in the panels (e) and (f) (Table 1). One likely explanation for this is that the internal noisiness of B and Y (their 'equivalent input noise', Pelli, 1990) is higher than the internal noisiness of R and G. Higher

internal noise levels in B and Y would cause their no-noise thresholds to be high compared to R and G, as they are, and would lessen the effect of adding external noise (because the external noise must overcome the internal noise to cause masking). Another factor is the probable nonlinearity of B and Y (see below); changing noise angle and contrast has unpredictable effects on nonlinear mechanisms.

3.2. Detection model

Threshold contours were fit to the detection data using a Minkowski summation rule (Graham, 1989; Quick, 1974):

$$\left(\sum_{i=1}^{K} |X_i|^{-C} \right)^{1/C} = 1 \qquad (1)$$

in which there are K mechanisms with responses X_i; the combination exponent C is described below. Each mechanism response is a half-wave rectified weighted sum of cone contrasts. For example, for the two classical pairs of mechanisms ($K = 4$):

$$X_1 = R = U_R \rho_1 \left(W_{R,L} \frac{\Delta L}{L} + W_{R,M} \frac{\Delta M}{M} + W_{R,S} \frac{\Delta S}{S} \right) \qquad (2a)$$

$$X_2 = G = U_G \rho_2 \left(W_{G,L} \frac{\Delta L}{L} + W_{G,M} \frac{\Delta M}{M} + W_{G,S} \frac{\Delta S}{S} \right), \qquad (2b)$$

$$X_3 = B = U_B \rho_3 \left(W_{B,L} \frac{\Delta L}{L} + W_{B,M} \frac{\Delta M}{M} + W_{B,S} \frac{\Delta S}{S} \right), \qquad (2c)$$

$$X_4 = Y = U_Y \rho_4 \left(W_{Y,L} \frac{\Delta L}{L} + W_{Y,M} \frac{\Delta M}{M} + W_{Y,S} \frac{\Delta S}{S} \right) \qquad (2d)$$

with the half-wave rectifiers defined so as to preserve the sign of the output:

$$\rho_i(x) = \begin{cases} x & (-1)^{i-1}x > 0 \\ 0 & (-1)^{i-1}x \leqslant 0 \end{cases}.$$

The U_i are mechanism sensitivities that may be interpreted in d' units (Eskew et al., 1999). Each vector of weights $\{W_{i,L}, W_{i,M}, W_{i,S}\}$ (the mechanism vector) is required to have unit length. Below we characterize these mechanisms using the sensitivity U_i and the angle made by the mechanism vector with $\{L_e, M_e, 0\}$. The 'red' and 'blue' polarities are defined to be positive.

Since the Gabor stimulus contains equal and symmetrically opposite contrasts, the thresholds must be symmetric about the origin. The symmetric nature of the detection results requires that the model mechanism vectors for each polarity (R and G, and Y and B) be negatives of one another; thus, their estimated gains (and their sensitivities under our conditions) are equal,

and their mechanism angles are separated by 180°. Such a symmetry is not required for the discrimination data, but because our discrimination model is based upon the detection model we effectively make a 'symmetry assumption' for the analysis of the discrimination data, requiring equal and opposite mechanism vectors for R and G, and Y and B. Experiments that have used unipolar stimuli provide excellent evidence in favor of the symmetry assumption for R and G, including the parallelism of the two detection contours (Cole et al., 1993; Cole, Stromeyer, & Kronauer, 1990; Giulianini & Eskew, 1998), and their equal distances from the origin, even under conditions of red and green chromatic adaptation (Chaparro, Stromeyer, Chen, & Kronauer, 1995; Eskew, Stromeyer, & Kronauer, 1992; Stromeyer et al., 1985). There is some evidence for symmetry between Y and B (e.g. Thornton & Pugh, 1983; Cole et al., 1993) and some against it (e.g., McLellan & Eskew, 2000; Shinomori, Spillmann, & Werner, 1999).

With the symmetry assumption, the sum of the appropriate pairs (R + G, and Y + B) is a linear, bipolar opponent channel cast as a function of cone contrasts. For the purposes of modeling the detection data, it would have been equivalent and simpler to use two bipolar processes (red\green and yellow\blue channels) rather than four half-wave rectified mechanisms (R, G, Y, and B mechanisms), but the analysis of the discrimination data requires having separate mechanisms for each chromatic polarity (Section 1).

The estimated relative L and M cone input to each mechanism is constrained by the choice of the equiluminant measurement plane, so there are two free parameters per mechanism (the mechanism vector length or sensitivity, and its angle). Symmetry means that there are no free parameters for half of the mechanisms (e.g., once the parameters for R are known, G's are also known). Thus, there are K free parameters in the model as applied here.

Each of the K linear mechanisms generates an isoresponse line in a plane of cone space, orthogonal to the mechanism vector. The combination rule of Eq. (1) rounds the corners of the convex polygon created by the set of such lines. A combination exponent of $C = 4.0$ was chosen on the basis of previous work (Cole et al., 1993; Cole, Hine, & McIlhagga, 1994; Eskew et al., 1999; see Appendix).

Eskew et al. (1999) reviewed several studies that estimated the cone contrast weights for the two classical pairs of detection mechanisms. The relative weights for the luminance, Y, and B detection mechanisms differ widely across observers and conditions, but there is hardly any variation for the R and G mechanisms (e.g. Eskew et al., 1999; Table 1). Thus, it is useful to use the standard observer weights we proposed for the R and G mechanisms ($W_{R,L} = -W_{G,L} = 0.70$, $W_{R,M} = -W_{G,M} = -0.72$, and $W_{R,S} = -W_{G,S} = 0.02$; see Eskew

et al., 1999; Table 2) to fix the R mechanism angle, increasing the degrees of freedom (as done by Giulianini & Eskew, 1998). With this additional constraint,

Table 2
Model Parameters with RG constrained at 1°\181°

Observer	Noise angle	R & G sensitivity	B angle	Y & B sensitivity
FG	0°\180°	30.2	63°	9.4
	120°\300°	20.2	56°	12.8
	90°\270°	96.8	3°[a]	110.0[a]
PK	0°\180°	35.0	50°	8.6
	145°\325°	31.8	52°	8.6
	90°\270°	71.8	6°[a]	53.8[a]
JRN	0°\180°	19.4	67°	13.0
JH	0°\180°	13.7	28°	12.0
RTE (Fig. 5)	0°\180°	20.7	71°	12.4.3
RTE (Fig. 7)	0°\180°	13.0	82°	8.6

[a] The fit produced an additional copy of the paired R and G mechanisms; there are no discernable Y and B mechanisms when the noise is along the S-cone direction.
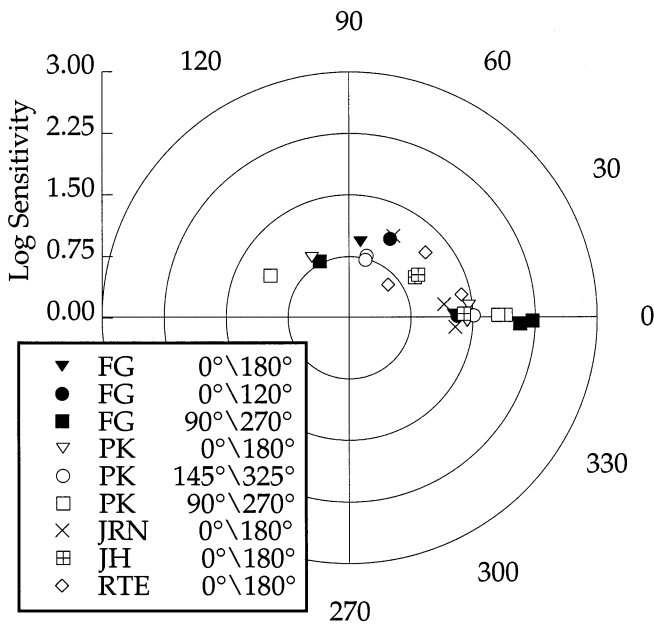

Fig. 2. A polar plot of the mechanism vectors resulting from the six mechanism model fits to the detection data of Figs. 1, 4–6 (note: the parameters of the standard model used to make discrimination predictions are given in Table 2, not here). Different symbols represent different subjects and noise color directions. The angular coordinate gives the mechanism vector direction, and the radial coordinate gives the $\log_{10}$ of the mechanism sensitivity (the vector length). Only the R and B mechanism vectors are represented; the G and Y ones are reflected across the origin. A cluster of points lies near the expected R direction (1°); the putative B vectors are much more scattered. Two of the three fitted mechanisms have essentially the same vector angle in every case, and the same sensitivity in every case but one (RTE, open diamonds).

there are $K - 1$ free parameters to be determined. For the $K = 4$, constrained R and G model, these free parameters are the R sensitivity, B sensitivity, and B angle; below we refer to this four-mechanism model (with the fixed R and G weights but variable R and G sensitivity) as the 'standard model.'

The solid lines in Figs. 1 and 4(a)–7(a) represent a best fit of the standard model to the masked detection data; the fits were excellent, with $R^2 > 0.97$ in all cases. The standard R mechanism vector $\{W_{R,L}, W_{R,M}, W_{R,S}\} = \{0.70, -0.72, 0.02\}$ has an angle of 1.14°–1.2° in these equiluminant planes, depending upon the observer's M_e/L_e ratio, illustrating again the tiny degree of angular variation produced by changes in the observers' equiluminant setting. For each of the ten data sets, the estimated sensitivities for the R and B mechanisms are given in Table 2, as estimated for the standard model. Due to symmetry, R and G, and Y and B, have the same sensitivity and mechanism vector angle. Note that when the noise was along the S cone 90°\270° direction, only R and G could be detected in the data.

For each set of masked detection data in Figs. 1(a)–(f) and 4(a)–6(a), we compared the standard model (three free parameters; results in Table 2), against two other models. The first of these is simply the best-fitting $K = 4$ model (R angle free to vary, so there are four free parameters). The second of these models allowed an additional pair of opponent mechanisms ($K = 6$), with all six parameters free to vary. We compared these models against the standard model using an F-statistic computed as the ratio of the two error χ^2 (each divided by its degrees of freedom). In all cases but one, we found that the standard, four-mechanism model could not be distinguished statistically from either of these alternatives: the slightly smaller error variances produced by the alternative models did not overcome their reduced degrees of freedom. The exception was for RTE (Fig. 5(a)), for whom the best-fitting R and G mechanisms had a larger S-cone input than the standard model shown ($F(10,9) = 5.67$, $P = 0.006$), and for whom the six-mechanism model performed better than the standard four mechanism one ($F(10,7) = 4.41$, $P = 0.02$). The smallest P-value for any of the other comparisons of the standard model versus $K = 6$ models was for Fig. 1(b), with $F(5,2) = 3.96$, $P = 0.21$. Therefore, with one exception, our detection data do not justify the use of additional mechanisms beyond the classical ones on statistical grounds, and the relative R and G weights proposed by Eskew et al. (1999) do as well as the best-fitting ones, as also found by Giulianini and Eskew (1998). Even for RTE (Fig. 5), the standard model performed better than the best-fitting model in predicting the discrimination data (next section).

To better illustrate what happens when a pair of nonclassical mechanisms is included in the model, Fig.

2 plots all of the $K = 6$ model fits (six free parameters). In this polar plot, each symbol represents one member of each of the three pairs of mechanism vectors from each set of masked data (i.e., R, B, and a nonclassical mechanism). The angular coordinate gives the direction of the mechanism vector, and the radial coordinate is the log of the mechanism sensitivity under the particular masked conditions (i.e., this is the masked sensitivity log U_i, not a parameter that is intrinsic to the mechanism). Four aspects of the figure deserve comment. First, many points lie near the 1° direction of the standard R mechanism (the mean angle is -0.1°), consistent with the good fits provided by the standard R and G mechanisms when the fits were so constrained (Figs. 1 and 3(a)–6(a)). Second, of the three mechanisms fit to each data set, two of them are essentially identical. For RTE (Fig. 5(a)), the only case in which the six-mechanism model was statistically superior to the standard four mechanism one, the two putative B mechanisms (open diamonds, Fig. 2) have the same vector angle (41°); they differ in sensitivity by a factor of four, however, so the weaker mechanism could not contribute to threshold (thus the statistical difference lies in the R and G weights, not the B and Y ones). This graphical analysis shows that even when the fitting algorithm was allowed to estimate three distinct pairs of opponent mechanisms, it did not do so, consistent with the statistical analysis reported above. Third, the angles for the second estimated mechanism – the putative B mechanism – differ rather widely, even within an observer, compared to the tightly clustered R directions. The standard error of each of these B angles (computed from the asymptotic standard errors of the fitted mechanism weights, using the appropriate propagation of error formula) was often quite large (up to 285°), whereas the R angle standard errors were generally smaller (most were less than 5°). Fourth, there is a wide range of masked R sensitivities, but within a condition the R mechanism is always more sensitive than the B mechanism, even when the noise was approximately aligned with the R and G mechanism directions.

The low B sensitivity is presumably one reason why the B direction estimates are so variable, but it also seems likely that the B (and Y) mechanisms are nonlinear, such that changes in noise characteristics could produce different approximations to the mechanism vectors. Most models of S-cone detection include strongly nonlinear elements (e.g. Pugh & Mollon, 1979; Zaidi, Shapiro, & Hood, 1992). Using noise masks that were seen only by the L and M cones, Giulianini (1998) recently showed that S cone threshold elevations could not be accounted for by any linear model, and McLellan and Eskew (2000) demonstrated different relative L and M inputs to the mechanisms responsible for detecting S-cone increments and decrements (a failure of symmetry as well as nonlinearities in B and Y).

In summary, the detection data of Figs. 1 and 4(a)–

7(a) provide no evidence for nonclassical chromatic detection mechanisms in the equiluminant plane, despite our efforts to use strong masking noise to reduce the sensitivity of the most-sensitive mechanisms and permit less-sensitive mechanisms to be revealed. A potentially more-informative method is a discrimination procedure used in conjunction with noise masking. By masking R and G mechanisms, we hoped to be able to reveal the presence of more than four labeled lines in the equiluminant plane.

3.3. Discrimination results

Fig. 3 will be used to illustrate the format of the data. Panel (a) shows the detection results as filled symbols when masking noise was used (and, in Figs. 4–7, as open symbols when noise was not used; JDP was only tested with noise). Under these noise conditions, JDP was unable to see the S-cone stimulus (Section 2), and thus we were unable to fit the detection model of Eq. (1) to his data. However, the lines have the appropriate slope for the standard R and G mechanisms (Eskew et al., 1999).

As described in Section 2, both the test and the standard were fixed at threshold contrast (82% 2AFC detection) for the discrimination task. Fig. 3(b) shows JDP's ability to discriminate threshold-level tests, in the presence of the 0°\180° masking noise, from a 180° standard stimulus. This polar plot has an angular coordinate representing the polar angle of the test color direction, and the distance of the filled square from the origin representing its discriminability from the standard color direction. The discrimination data need not be symmetric about the origin, and are only plotted once (at the angle corresponding to the chromaticity of the bar of the Gabor just above fixation). The circumference of the gray disk corresponds to chance-level discrimination performance. The figure shows that JDP could not discriminate any of the four most extreme points we measured (at 105°, 112°, 240°, and 255°) from the 180° test direction – discrimination was always at chance. Under these masked conditions, JDP's thresholds reflect only R and G mechanisms.

Similar data are shown in Figs. 4–7 for three additional observers. The arrowheads in each panel show the 'corner' angle of the *detection* contour, determined by the intersection of the straight line threshold lines for the mechanisms fit to the detection data (i.e., the lines of Eq. (2a)–(2d) without the Minkowski combination of Eq. (1)). These arrowheads are provided to make comparisons across panels easier; note, however, that the corner angles in the detection contour should not necessarily correspond exactly to those stimuli that are imperfectly discriminated from the standards (Section 4). In panels (b)–(d), the small open squares show model predictions that are described in the next section.
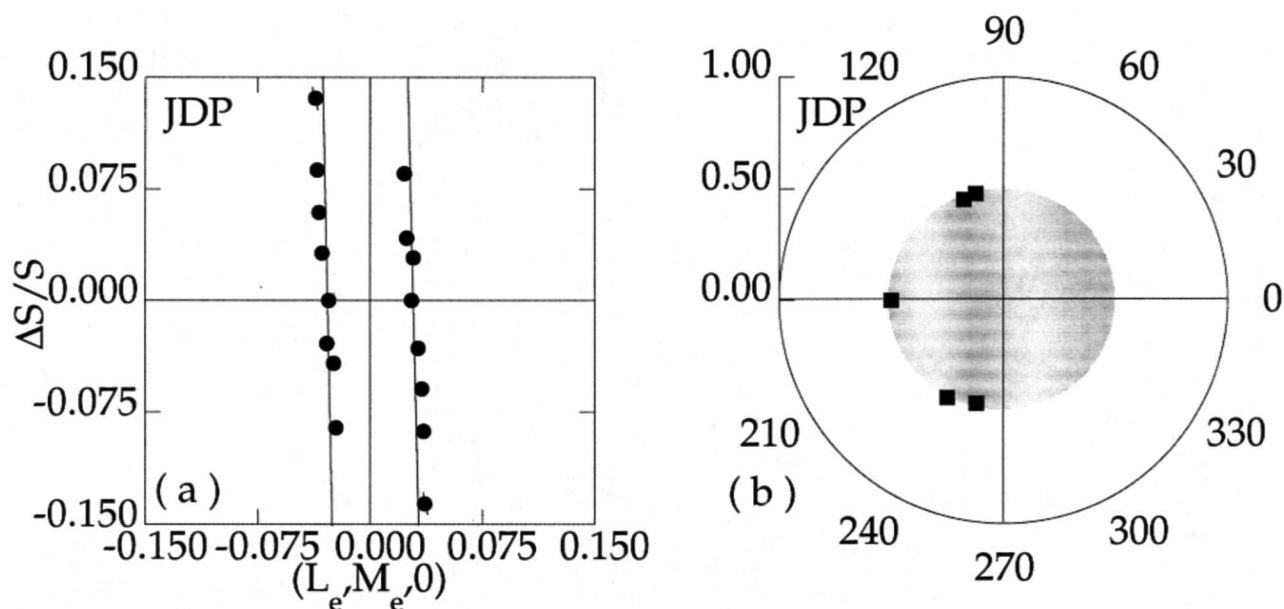
Fig. 3. Detection and discrimination, observer JDP (190 Td, noise contrast = 0.11). (a) Detection contour, as in Fig. 1. The solid lines are drawn by eye, with a slope appropriate to the standard R and G detection mechanisms from Eskew et al. (1999) – the small S-cone input to R and G mechanisms determines the slope of the line. S-cone thresholds were not measurable for this observer under these conditions. Error bars are as in Fig. 1. (b) Discrimination data, in polar coordinates. The angular coordinate is the test color angle, and the radial coordinate is discriminability from the 180° standard, with chance (50%) discriminability at the circumference of the gray disk. For clarity, a filled square has been added at 50% discriminability at 180°, although we did not actually measure the discriminability of the standard against itself.

JRN and RTE (Figs. 4 and 5) were tested most extensively; we varied the test direction in small steps near the corner of the detection contour, to attempt to find a region that was well discriminated from both the S and L\M direction, which would indicate that we had isolated an additional mechanism. Fig. 4(b) shows discriminability against the 90° ($+\Delta S/S$) standard for JRN. The tests at 68° through 82° could not be discriminated from 90°, whereas the 45° test is well discriminable from 90°. Conversely, Fig. 4(c) shows that the tests at 68°–90° are as highly discriminable from 0°, whereas the 45° test cannot be discriminated from the 0° direction at all. Thus from about 68° to 90° is a spectral band that represents a labeled line mechanism.

As shown in Fig. 4(b) and (c), the 52° and 60° tests, which are near the corner of the detection contour (arrowheads), are imperfectly discriminable from either the 0° or 90° standards (at about 60%), consistent with these stimuli being detected by the probability sum of two mechanisms. On some trials one mechanism detects these tests, on other trials the other mechanism detects these tests, and on some trials both mechanisms do. On those trials in which both the test and standard are detected by the same mechanism, the observer cannot discriminate them, but on the other trials the observer may be able to do so, resulting in an intermediate level of performance. This good correspondence between the corner angles of the detection contour and these imperfectly discriminable test angles is due to the good fit of the *detection* model to the thresholds in panel (a) (Section 4).

The open diamond symbols in both Fig. 4(a) and (c) were obtained in a later, separate experiment, and were not included in the model fits to the detection data. They were collected to be certain that none of the angles below 45° were well discriminated from 0°.

Discrimination against the 270° standard is shown in Fig. 4(d). The 278° and 285° tests are completely indiscriminable from 270°. The 292° and 300° tests are imperfectly discriminated from 270° (63% and 72% correct, respectively). Again these transitional stimuli correspond well to a corner of the detection contour. The 315° test was well discriminated from the 270° standard.

Two Gabor stimuli separated by 180° in the plane are identical except for a 180° spatial phase shift in their carrier sine waves, and the results indicate that the observers generally were able to use this spatial phase information to make discriminations. The 180° test was highly discriminable from the 0° standard (Fig. 4(c)), and the 270° test is highly discriminable from the 90° standard (Fig. 4(b)).

Using the same logic as for JRN, the data in Fig. 5(b) and (c) show that for RTE there are spectral bands from about 135° to about 225°, and from about 82.5° to 97.5°, that correspond to labeled line mechanisms. The corners of the detection contour (arrowheads at 51° and 118°) correspond fairly well, but not perfectly, to the stimuli that are imperfectly discriminable from both 180° and 90° standards, reflecting the imperfect fit of the detection model in panel (a) (Section 4).

Similar results were obtained from observer JH (Fig. 6). For this observer also, the data indicate the presence of only two labeled line mechanisms in a given quadrant of the equiluminant plane. Here, the correspondence between the arrowheads and the transitional discrimination performance is not very good (Section 4).

For RTE, Fig. 5(d) shows discrimination against a 115° standard, selected to be near the corner of the detection contour. Discrimination goes from chance up

to ~ 60% quickly on either side, unlike the broader bands seen when the standard was at 180° or 90°, and discriminability is still intermediate with both the 90° and 180° tests, consistent with 115° being approximately equally detected by two mechanisms.

Comparing panels (b) and (c) within Figs. 4–6 (and panel (d) in Figs. 4 and 6) indicates good agreement, within a subject, of the angles at which intermediate-level discrimination performance was obtained. This consistency of transition angles across different stan-
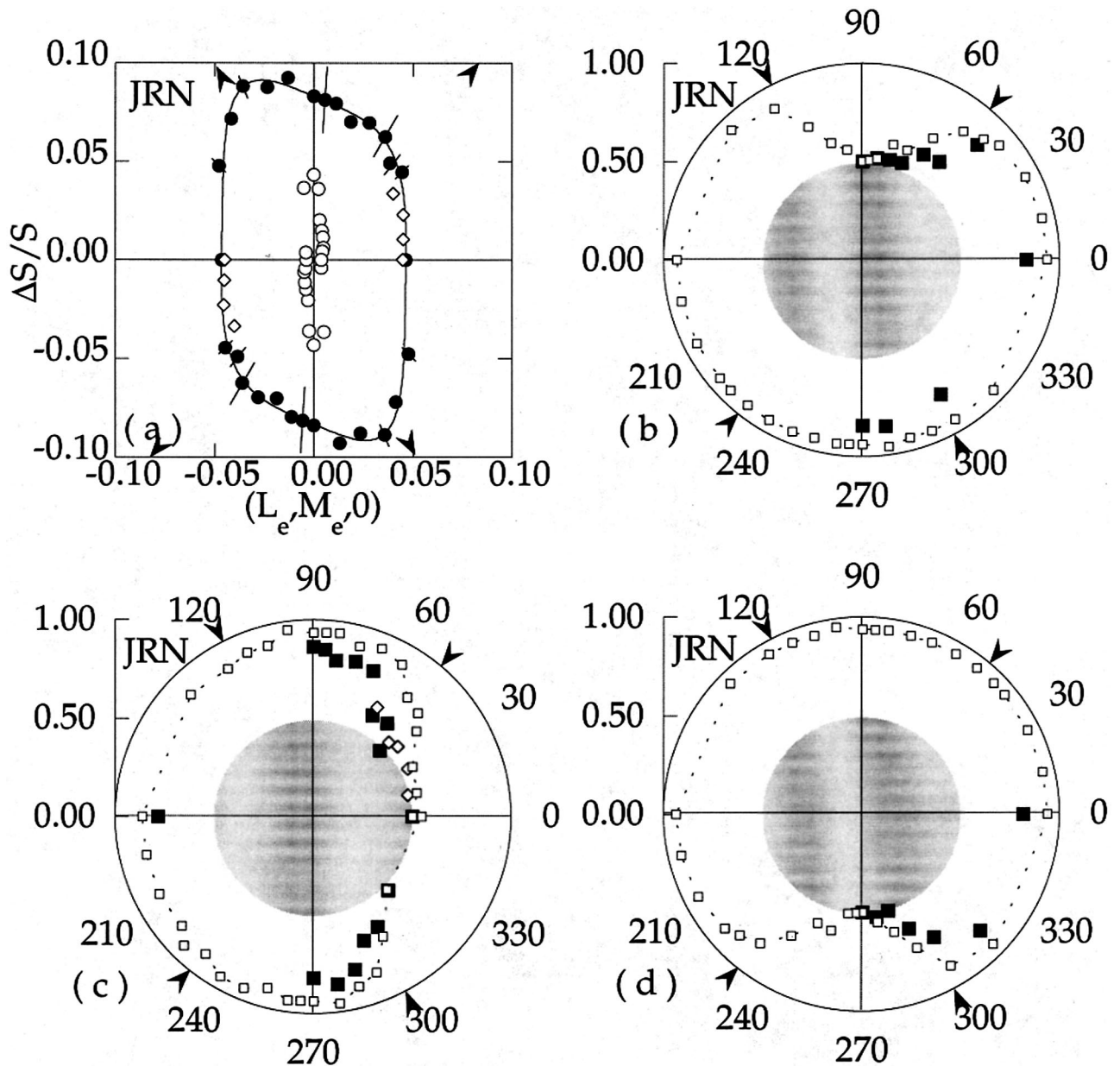

Fig. 4. Observer JRN (190 Td, noise contrast = 0.10). (a) Detection contour. The solid line is the standard model fit and the error bars are as in Fig. 1. Panels (b), (c), and (d) represent discrimination data, with standard angles of (b) 90°, (c) 0°, and (d) 270°. In each panel, the angular coordinate is the test color angle, and the radial coordinate is discriminability from the fixed standard, with chance (50%) discriminability at the circumference of the gray disk. For clarity, a filled square has been added at 50% discriminability at the angle corresponding to the standard itself, although we did not actually measure the discriminability of the standard against itself. To prevent clutter, error bars are not plotted, but these bars are almost always less than twice the size of the symbols. Open squares give the prediction of the Bayesian classifier (refer text).
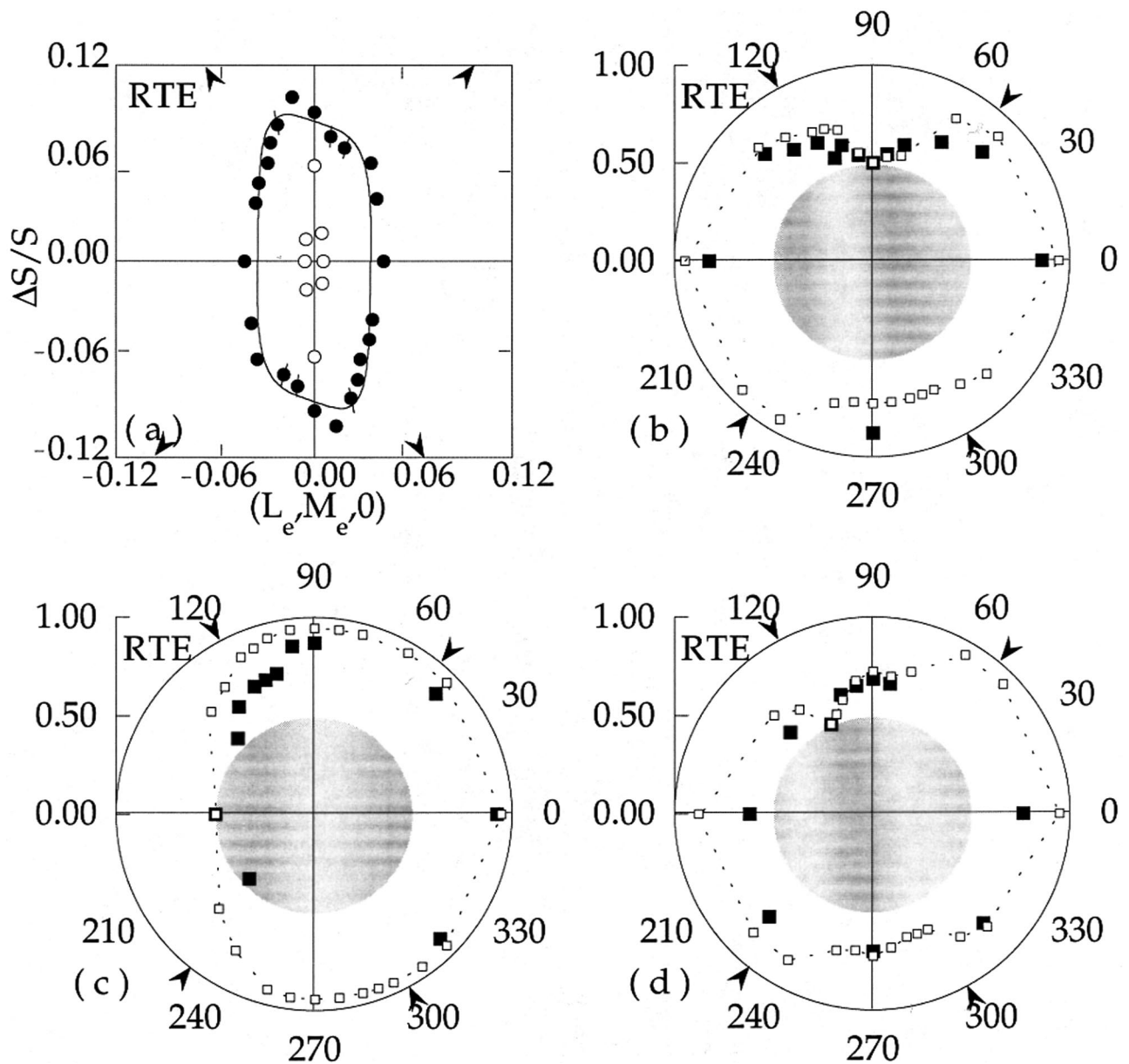
Fig. 5. As in Fig. 4, but for observer RTE (190 Td, noise contrast = 0.06). (a) Detection contour. Panels (b), (c), and (d) represent the discrimination data, with standard angles of (b) 90°, (c) 180°, and (d) 115°.

dards suggests that there are fixed color boundaries, rather than the discrimination being based upon a distance in color space. The results in Fig. 5(d) with the intermediate standard also show that the transition angle is fixed, independent of the standard angle. Fixed boundaries of this sort have been interpreted as evidence for categorical color perception (Mullen & Kulikowski, 1990; Wandell, 1985), which in turn implies a limited number of color mechanisms.

As shown in panel (a) of Figs. 4–6, the noise raises the thresholds for the R and G-detected tests near 0°\180° by factors of about 12, 8, and 5 for JRN, RTE, and JH, respectively. Thus, there is ample opportunity for less-sensitive nonclassical mechanisms to be revealed in the

discrimination data. However, to ensure that we sufficiently masked the R and G mechanisms and explored the top and bottom parts of the detection contour, we repeated the experiment with higher contrast noise for RTE. As shown in Fig. 7(a), with this noise we did not have sufficient test contrast in our apparatus to measure the complete detection contour. Fig. 7(b) shows an indiscriminability band centered on the 90° direction (∼ 70°–105°) which is somewhat broader than the band obtained with weaker noise (Fig. 5(b)). About five points could be measured on the apparent B threshold line with the stronger noise; at no test angle was there evidence of improved discriminability, which would have indicated intrusion by an additional mechanism.

In summary, the chance-level discrimination performance over the middle portion of the plane when using the red/green standard is inconsistent with the presence of labeled line mechanisms besides R and G in this region. Additional data, not shown, show an even more R and G-dominated pattern of discrimination without the masking noise. Narrower bands around 90° and 270° indicate two additional labeled line mechanisms, presumably Y and B. It is clear that, despite masking the R and G to the greatest degree possible with our apparatus, we were unable to isolate any additional detection mechanisms beyond the classical ones.

3.4. Discrimination model

The open squares in the discrimination plots show the performance of a Bayesian classifier developed for this task. The inputs to the model are the color angles and vector lengths representing the test and standard stimuli (defined as the Gabor peak just above fixation). The model classifier chooses between test ('I') and standard ('II') stimuli based upon the posterior probability

respond   "I"   if $P(\text{I} \mid)\Omega \geqslant P(\text{II} \mid)\Omega$
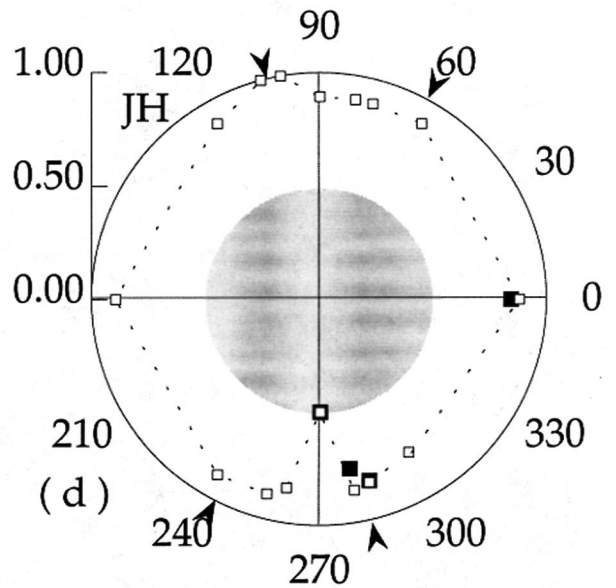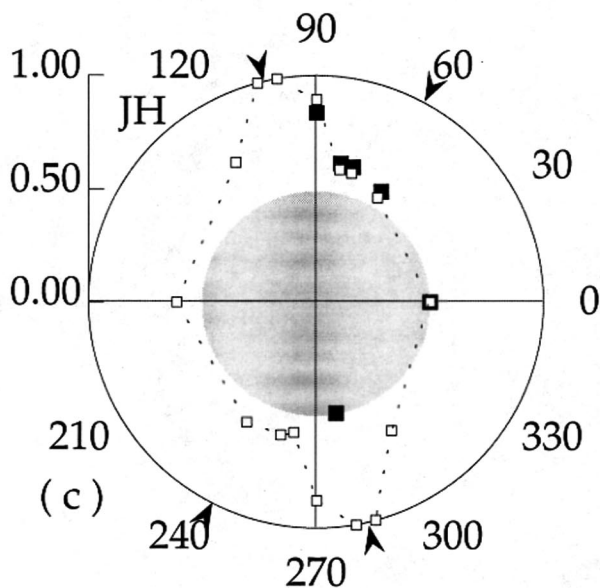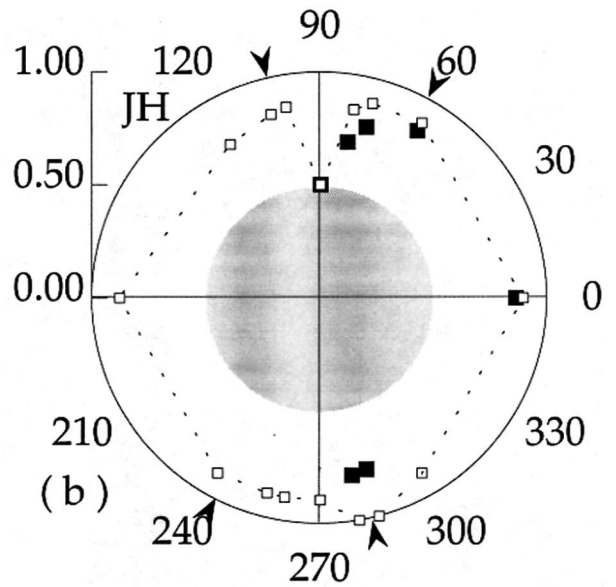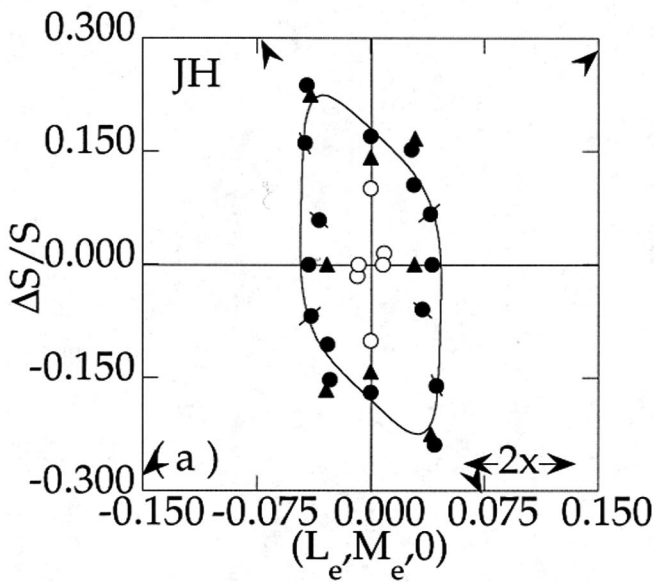   "II"   otherwise

Fig. 6. As in Fig. 4, but for observer JH (190 Td, noise contrast = 0.08). (a) Detection contour. Triangles are replications, which were included in fitting the model. Panels (b), (c), and (d) represent discrimination data, with standard angles of (b) 90°, (c) 0°, and (d) 270°. The filled circles in panel (a) represent the stimuli that were used in the discrimination experiment.
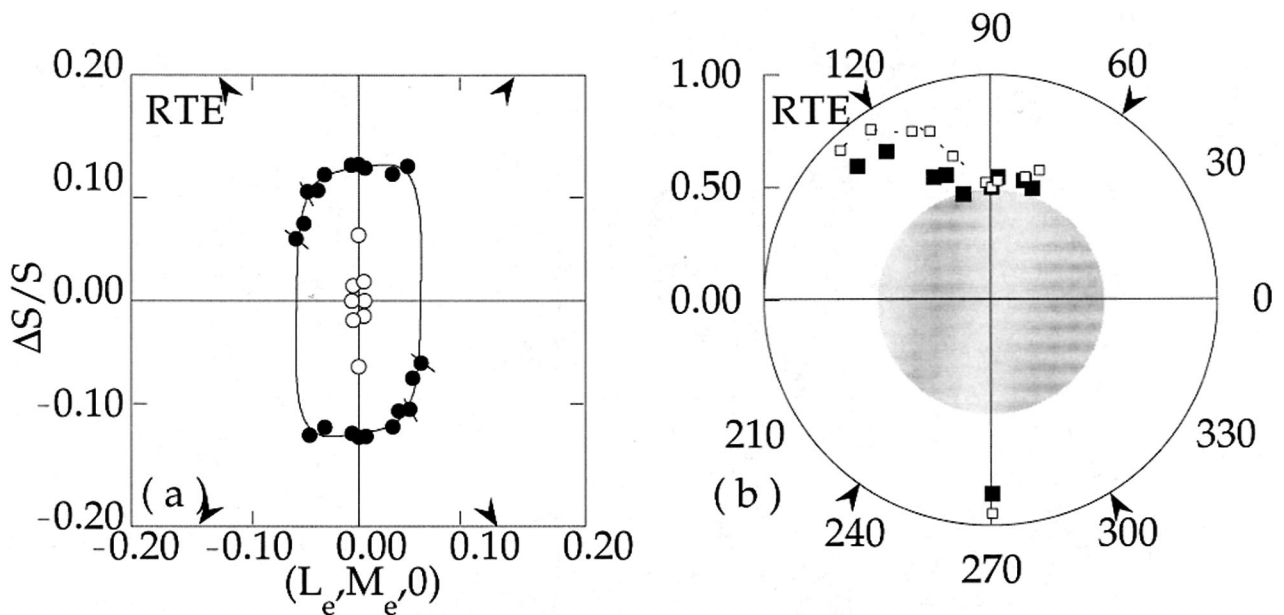
Fig. 7. As in Fig. 4, but for observer RTE (190 Td, noise contrast = 0.09). (a) Detection contour. With the stronger noise (compared to Fig. 5), we did not have enough test contrast to measure but a few thresholds on the vertical segments of the detection contour. Panel (b) represents discrimination data, with a standard angle of 90°.

with Ω_i representing the set of binary decisions made by each mechanism during a single stimulus presentation. For example, suppose that in interval 1 of a given trial, the R and B mechanisms of Eq. (2a)–(2d) detect the stimulus, but the other two do not; then $\Omega_i = r\bar{g}b\bar{y}$ (evidence for a 'purple' stimulus). The classifier receives a set of mechanism detections Ω, and makes the optimal decision based upon those mechanism outputs by choosing the stimulus with the greatest likelihood. Details of the model are provided in Appendix A.

For each of the mechanisms of Eq. (2a)–(2d), the mechanism vector angle and sensitivity were taken from the detection model fits shown in Table 2, and used to predict the detection responses of each of the four mechanisms to test and standard stimuli. These responses were then optimally combined, according to the Bayesian classifier (Appendix A) and used to generate discrimination probabilities, shown as the open squares in Figs. 4–7. Note that there were no free parameters in these model predictions. Although the Bayesian classifier generally performs better than the human observers, the pattern of predicted performance is very similar to the data. For example, in Fig. 4(b), the model predicts that discriminability against the 90° standard should rise gradually as the test angle is lowered from 90°, to reach an asymptote near 50°; the data follow this pattern quite well. Fig. 4(c) shows the model performance with a 0° standard; again the pattern of the actual data is well predicted, although the model does better than the human observer, especially near 50°. Note that the model predicts very high discriminability of the 180° test from the 0° standard, two stimuli which

are 180° spatial phase shifts of one another. There is no explicit phase sensitivity in the model, but because the input to the model is the Gabor peak just above fixation and this peak is 'red' at 0° and 'green' at 180°, the model correctly predicts good discrimination in this case. The single biggest discrepancy between prediction and data is in Fig. 6(c), in which JH was completely unable to discriminate the 280° test from the 0° standard; this may have been due to the failure of JH to attend to the phase of the B\Y component of the Gabor.

In one of RTE's detection contours (Fig. 5(a)), the best-fitting four mechanism model has a larger S-cone input to R and G than the standard model that is plotted (the best-fitting contour has vertical segments with larger tilts). Yet when the best-fitting detection mechanisms were used as the basis for the discrimination model, the predicted discriminabilities were somewhat less accurate than those based upon the standard model (open squares, Fig. 5(b)–(d)). This may indicate that RTE's R and G cone weights are actually more similar to the standard model's than the best-fitting ones despite the statistical result to the contrary (Section 3.1); however, we did collect some additiona data to check and these thresholds were consistent with the original ones, so the issue remains unresolved.

Although the discrimination model is not, strictly, an ideal observer – it is an ideal classifier of the binary inputs provided by the sub-optimal chromatic mechanisms – it is not surprising that it would do better than a noisy human decision maker. Both the model and the human are limited by stimulus information at the upper

end (and in any case performance cannot exceed 100%), and of course neither model nor human can do worse than chance. The implication is that the model should do better than the observers mainly at intermediate levels of discriminability, and this is in general what happens: as the angular difference between test and standard stimuli increases, the predictions rise more rapidly from chance than the actual data do, then approach an asymptotic level of performance that is slightly higher than the actual discriminability. Because of this overprediction of performance at intermediate discriminability – a curvilinearity in the scatterplots of actual vs. predicted discriminability – the R^2 values are not high (most are about 0.4), but nonetheless the model provides a very good description of the pattern of discrimination performance, with no free parameters.

We have not exhaustively explored other possible models, largely because there is no agreed-upon, computable model of nonclassical color mechanisms. However, adding additional (nonclassical) mechanisms to our model would in general cause it to predict better performance than the four mechanism model does. Consider the 'purplish' tests near 45° in Fig. 4(c), for example. If there were a narrowly tuned mechanism that contributed significantly to detection of these test stimuli, but not to detection of the 0° standard, then the model would predict *better* discriminability, based upon those trials in which the test was seen only by the 'purple' mechanism. This would increase the discrepancy between the model and the actual data in this region. Similar considerations apply elsewhere: as a rule, the model fit with $K > 4$ would be worse than it is with four mechanisms.

4. General discussion

The detection contours are well described by a combination of four mechanisms, and attempting to force the existence of an additional pair of detection mechanisms fails (Fig. 2). The discrimination data of Figs. 3–7, interpreted using standard assumptions about labeled-line detectors, suggest that at most only four mechanisms contribute to seeing color under our conditions. Despite the masking of R and G, no stimulus was predominantly detected by a nonclassical labeled line mechanism: there was no intermediate band of improved discriminability that would indicate the presence of such a mechanism. This conclusion does not depend upon the details of either the detection or discrimination models.

The Bayesian classifier model provides a very good description of the pattern of discriminability, with no free parameters. In general the discrimination model performs better than the human observers, as it should. By definition, a labeled-line mechanism can discriminate stimuli as well as it can detect them (Introduction, Watson & Robson, 1981). Our model obeys this principle, but the obtained performance level necessarily depends upon the psychophysical task. In the detection task there was one stimulus presentation per trial, whereas there were two presentations in each trial of the forced-choice discrimination task, each one of which suffices to correctly make the discrimination. After correcting for the difference between a single presentation (in the detection task) and the double presentation (in the discrimination task), the best performance of our model in discriminating two stimuli that are exactly at threshold is 0.86 (Appendix A).

A related point has to do with how the contrasts of the test and standard affect the relationship between the detection and discrimination models. In Figs. 4 and 5, the corners of the detection contours (arrowheads) correspond well to those stimuli that are discriminated at an intermediate level ($\sim 60\%$–70%), indicating detection by multiple mechanisms. In Figs. 6 and 7, the corners agree less well with the transitional discrimination angles. This is not a failure of the discrimination model to fit the data: note that the *predicted* transitional discriminations in Figs. 6 and 7 (open squares) also deviate slightly from the corner angles. The lack of agreement with the arrowheads reflects instead small deviations of the thresholds from the *detection* model used to determine the corners. In the discrimination experiment we set the standard and test stimuli to their measured threshold contrast levels, which do not lie exactly on the fitted detection contours. If, for example, the standard stimulus is below the threshold contour (Fig. 6(a), 90°), and the test stimulus is above it (Fig. 6(a), 80°), then, the discrimination model will predict better discriminability due to their contrast differences (Fig. 6(b)). To the extent that the detection model — the continuous lines in Figs. 4–7(a) — better represents the true 82% detection of the observer than do the noisy individual threshold estimates, the observer's performance will reflect this contrast cue, as does the model. The net result of these contrast variations – either above or below threshold – in the test and standard stimuli is that discrimination performance need not be intermediate at exactly the corner angle of the *detection* model. The discrimination model predictions (open squares, Figs. 4–7) take this contrast deviation into account, but the arrowheads do not. Overall, the better the fit to the detection data, the better the discrimination model should describe the discrimination data, and the closer the correspondence should be between detection and discrimination models.

An important general point to be made here is that complex patterns of detection and discrimination results can be generated by classical mechanisms alone. An example may be seen in the inflections in the behavior of the discrimination model with the 0° stan-

dard (e.g., open squares near 300° for JRN, Fig. 4(c)). These abrupt changes in performance come about because the mechanisms are not orthogonal, and the psychometric functions are not linear. Under these circumstances, the classical four-mechanism model can produce behavior that might have been incorrectly interpreted as indicating the presence of additional mechanisms had it been seen in isolation. An analogous point can be made with regard to the recent results of Stromeyer, Thabet, Chaparro, and Kronauer (1999), who showed that when a periodic test, such as a grating patch, stimulates the paired red and green mechanisms and the luminance mechanism approximately equally, a spatial phase interaction can distort the shape of a detection contour, producing the appearance of nonclassical chromatic mechanisms even when they do not exist. Demonstrating inflections in detection or discrimination contours is not sufficient to prove that nonclassical mechanisms are at work.

When the S cones are unmodulated, most detection experiments have found evidence for only the classical chromatic mechanisms (e.g. Giulianini & Eskew, 1998; Sankeralli & Mullen, 1996, 1997; Stromeyer et al., 1999), whereas many detection experiments in the equiluminant plane where S cones are modulated have concluded there are additional mechanisms (e.g., D'Zmura, 1991; D'Zmura & Knoblauch, 1998; Krauskopf et al., 1986; Zaidi & Halevy, 1993). The present experiments, like those of Sankeralli and Mullen, found no hint of additional mechanisms despite using both noise and test directions that stimulated S cones. There are several possible reasons for the discrepancy: there could be genuine individual differences (perhaps our observers are 'lower order'), and there are procedural and experimental differences between the studies, especially with regard to the size and retinal location of the test stimuli. For example, D'Zmura and Knoblauch used a Gaussian blob with $\sigma = 2.4°$ and Zaidi and Halevy used an 8° spot, centered on the fovea; the search task of D'Zmura used small spots scattered over a $7.2° \times 7.2°$ region; and the color matching task of Webster and Mollon (1994) employed two 2° squares located just outside the central fovea. Our relatively brief, foveated 1 cpd Gabor tests, as well as the 1 cpd patterns used by Sankeralli and Mullen, might have produced less activity by nonclassical mechanisms fed by S cones. Parametric variation in stimulus size and retinal location might be quite informative.

Color discrimination experiments have also produced discrepant results. Krauskopf et al. (1986) found evidence for $K > 4$ (more than the two classical opponent channels) in the equiluminant plane using a discrimination task. However, using long-wavelength lights that did not stimulate S cones, Calkins, Thornton, and Pugh (1992) showed that increments presented on a 578 nm background produced three spectral bands, consistent with red, incremental luminance, and green mechanisms only. Also using an increment threshold procedure, but with a white background and tests that varied over the entire spectrum, Mullen and Kulikowski (1990) found four perfectly discriminable regions at threshold, labeled 'orange', 'yellow', 'green', and 'blue,' and a fifth region ('violet') that was imperfectly discriminable from the blue region. The imperfect discriminability of these short-wave tests is consistent with there being an S-cone input into the red and green detection mechanisms (Eskew & Kortick, 1994; Ingling, 1977; Stromeyer et al., 1998; Werner & Wooten, 1979), with increment S-cone signals causing both redness and blueness. The present detection and discrimination results also indicate a small S-cone input into R and G mechanisms.

Although the discrimination results of Mullen and Kulikowski (1990) suggest that there are only four chromatic mechanisms, nonclassical mechanisms might be less sensitive, and, if so, discrimination tasks performed near threshold would not reveal them. The results of Calkins et al. (1992) implying only red, green, and luminance mechanisms held for stimuli as much as five times detection threshold, but their stimuli did not modulate S cones. Our stimuli were detectable by S cones, and the masking noise substantially raised the threshold of the red and green mechanisms, yet we found no evidence for discrimination by nonclassical mechanisms.

No one disputes that people can discriminate suprathreshold intermediate hues, for example an orange, from unique hues such as red. Thus at some level of the nervous system there surely must be a cell or cells that respond best to intermediate hues (and indeed there seem to be: Gegenfurtner et al., 1997; Kiper et al., 1997; Lennie et al., 1990). The question is how such cells can be isolated by psychophysical methods, and, once isolated, how they may be enumerated and characterized. In our laboratory we have yet to find an experimental procedure that reveals nonclassical psychophysical color mechanisms, rigorously defined, in the parathreshold regime. Perhaps these nonclassical cells are insensitive under our conditions, or perhaps they do not behave such as to satisfy the criteria we use to define psychophysical mechanisms; for example, perhaps they are not labeled lines.

Appendix A. Bayesian color classifier

In the discrimination procedure used here, each trial consisted of two intervals, one containing a standard stimulus and the other a test stimulus, presented in random order, with both stimuli fixed at detection threshold. The observer could be correct on any trial by correctly classifying the stimulus in interval 1 (with probability P_1), interval 2 (with probability P_2), or both; if the observer was not able to classify the stimulus in either interval, we assume that he or she guessed (with probability correct 1/2). The two intervals are experimentally symmetric, and the treatment of each interval includes effects of both the test and the standard (see below). Therefore $P_1 = P_2 = P$, and we assume that the intervals are independent, so that the probability of not correctly classifying the stimulus in either interval is $(1 - P)^2$. Therefore, the overall probability correct is given by the complement of not guessing correctly and not classifying correctly:

$$P_{\text{tot}} = 1 - \left(1 - \frac{1}{2}\right)(1 - P)^2. \tag{A1}$$

Denote the set of possible binary mechanism decisions by Ω, which has power 2^K. This set contains all the possible response patterns with each mechanism giving a 'detect' or 'nondetect' output. For example, for the four mechanisms given in Eq. (2a)–(2d),

$$\Omega = \{rgby, \bar{r}gby, r\bar{g}by, rgb\bar{y}, rg\bar{b}y, \bar{r}\bar{g}by, \bar{r}g\bar{b}y, \ldots, \bar{r}\bar{g}\bar{b}\bar{y}\}$$

with the first element representing 'true' detection by all four mechanisms, the second element representing nondetection by R and detection by the other three, etc. For these four mechanisms there are sixteen elements (compound responses) in Ω.

Denote the test and standard stimuli with roman numerals I and II, respectively. On a given interval of a trial, the mechanism response pattern is Ω_i (for example, on a given trial the observer's visual system might respond $r\bar{g}y\bar{b}$, evidence for an 'orange' stimulus being presented). Because the prior probabilities are equal, $P(\text{I}) = P(\text{II}) = 1/2$, the ideal classifier of a single stimulus bases its decision on the posterior probability:

respond "I" if $P(\text{I} \mid)\Omega \geqslant P(\text{II} \mid)\Omega$ (A2)

  "II" otherwise

This is a maximum likelihood classifier (see, for example, Kay, 1998, Chapter 3).

The probability that the decision specified by (Eq. (A2)) is correct, excluding guesses, is given by the absolute difference $|P(\text{I}|\Omega_i) - P(\text{II}|\Omega_i)|$. To obtain the overall 'true decision' performance of this classifier, we weight each of these probabilities of being correct by the probability that they occur, and sum over all the response patterns in Ω:

$$P = \sum_{i=1}^{2^K} |P(\text{I}|\Omega_i) - P(\text{II}|\Omega_i)| P(\Omega_i). \tag{A3}$$

The probability that a given response pattern Ω_i will occur is

$$P(\Omega_i) = P(\Omega_i|\text{I})P(\text{I}) + P(\Omega_i|\text{II})P(\text{II})$$
$$= \tfrac{1}{2}[P(\Omega_i|\text{I}) + P(\Omega_i|\text{II})].$$

To apply the model, we must first determine the compound posterior probabilities using Bayes' rule:

$$P(\text{I}|\Omega_i) = \frac{P(\Omega_i|\text{I})}{P(\Omega_i|\text{I}) + P(\Omega_i|\text{II})} \tag{A4}$$

since the priors are equal. We assume that the mechanisms are independent, so that the compound probabilities on the right-hand side of (Eq. (A4)) are given by the product of the simple probabilities. For Ω_7 from the above example, $P(\Omega_7|\text{I}) = P(\bar{r}|\text{I})P(g|\text{I})P(\bar{b}|\text{I})P(y|\text{I})$. Note that we assume the R and G mechanisms are stochastically independent – the noises in each mechanism are uncorrelated – but because of the symmetry and high-threshold assumptions the stimulus can never produce a response in both polarities of a channel simultaneously (so, $P(r|\text{I})P(g|\text{I}) = 0$ for example). These assumptions imply that many elements of Ω have zero probability, but the general model need not include the symmetry assumption.

The probabilities of mechanism detection conditional upon stimulus ($P(r|\text{I})$, etc.) are determinable from the model fits to the detection thresholds, with two assumptions. First, we have to determine the actual performance level that is relevant for the single interval procedure, and second, we must make assumptions about the psychometric slope.

As an initial step towards the first issue, we must define the test and standard stimuli used as inputs to the discrimination model, and there are two possible ways to do so. We could use the actual threshold values determined in the detection procedure (the actual stimuli presented in the discrimination experiment). Alternatively, we could fit the detection model and use the model to *predict* the threshold values. The first method uses the data points in panel (a) of Figs. 4–7 as the model inputs. The second method uses points on the continuous lines in panel (a) of those figures as the model inputs. A case can be made for either method, but we have used the first one. Using the second approach produces more accurate predictions in some cases (especially for JRN) and worse ones in others (especially for JH); for all the observers the second approach would cause the model predictions of transitional discriminability to correspond more closely to the arrowheads that mark the detection contour contours (Section 4). To the extent the detection model fits the thresholds, the two approaches are equivalent.

Next we must determine the relationship between the two-interval, single-presentation detection procedure

and the model analysis of the discrimination task (which is necessarily based upon the analysis of a single stimulus presentation; two such presentations occur in each trial). A stimulus for which detectability was 0.82, the level used to define threshold in the 2AFC procedure, should be less detectable in a single interval. In general, d' in a 2AFC procedure is $\sqrt{2}$ larger than d' in a single-interval yes/no procedure (e.g., Green & Swets, 1974), and this relationship was used here to define the threshold detectability in the application of the model. Thus, $d'_{2AFC} = 1.29$ corresponds to $d'_{YN} = 0.91$, or a probability of 2AFC detection of 0.74. This last value was 'corrected for guessing' to 0.48, to give the 'true' probability of detection at single-interval detection threshold by any mechanism. For example, we assumed that where the detection model indicated a stimulus was detected by the R mechanism alone, $P(r|I) = 0.48$, and $P(\bar{r}|I) = 0.52$. This $\sqrt{2}$ correction sets the highest level of discriminability achieved by the model. With the $\sqrt{2}$ correction, the model discriminates between two stimuli that are detectable exactly 82% of the time (according to the *detection* model) at 0.86; the predicted discriminability in Figs. 4–7 go up as high as 1.00 because the test or standard or both were not exactly at 82% detection threshold according to the *detection* model (Section 4).

Last we must make assumptions about the shape of the psychometric function for detection. The slope of the Weibull psychometric function for R and G detection mechanisms is ≈ 2.0, at least when relatively low spatial and temporal frequencies predominate in the test (Eskew, Stromeyer, & Kronauer, 1994; Eskew, Stromeyer, Picotte, & Kronauer, 1991), and Y and B mechanisms are similar (Watanabe, Smith, & Pokorny, 1997). However, using 2.0 as the exponent in the Minkowski combination of Eq. (1) to model the detection data would have produced poor results: as noted previously (Cole et al., 1993; Eskew et al., 1999), measured chromatic detection data are too 'squared off' in the corners to be consistent with a probability sum of mechanisms with slope two (which would be ellipses). Therefore, we used a value of four for the detection combination exponent and, for consistency, as the assumed psychometric slope in the discrimination model. Using a lower slope improves the fit of the model to the discrimination data, but only slightly.

In this model we used the high-threshold assumption to make corrections for guessing, although high-threshold theory is clearly wrong (e.g. Graham, 1989). We regard these corrections as approximations to the correct analysis, which would use the decision variable and criterion of signal detection theory. The fundamental elements of our discrimination model, including the combination of binary chromatic mechanism decisions, can be compatible with signal detection theory (see, for example, Pelli (1985) for related discussion).

References

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. New York: Irvington.

Boynton, R. M. (1978). Ten years of research with the minimally distinct border. In J. C. Armington, J. Krauskopf, & B. R. Wooten, *Visual psychophysics and physiology* (pp. 193–207). New York: Academic Press.

Boynton, R. M., Nagy, A. L., & Olson, C. X. (1983). A flaw in equations for predicting chromatic differences. *Color Research and Application*, 8, 69–74.

Calkins, D. J. (1999). Synaptic organization of cone pathways in the primate retina. In K. Gegenfurtner, & L. T. Sharpe, *Color vision: from genes to perception* (pp. 163–179). Cambridge: Cambridge University Press.

Calkins, D. J., Thornton, J. E., & Pugh, E. N., Jr. (1992). Monochromatism determined at a long-wavelength/middle-wavelength cone-antagonist locus. *Vision Research*, 32, 2349–2367.

Chaparro, A., Stromeyer, C. F., III, Chen, G., & Kronauer, R. E. (1995). Human cones appear to adapt at low light levels: measurements on the red-green detection mechanism. *Vision Research*, 35, 3103–3118.

Chaparro, A., Stromeyer, C. F., III, Huang, E. P., Kronauer, R. E., & Eskew, R. T., Jr. (1993). Colour is what the eye sees best. *Nature*, 361, 348–350.

Cole, G. R., Hine, T., & McIlhagga, W. (1993). Detection mechanisms in L-, M-, and S-cone contrast space. *Journal of the Optical Society of America A*, 10, 38–51.

Cole, G. R., Hine, T. J., & McIlhagga, W. (1994). Estimation of linear detection mechanisms for stimuli of medium spatial frequency. *Vision Research*, 34, 1267–1278.

Cole, G. R., Stromeyer, C. F., III, & Kronauer, R. E. (1990). Visual interactions with luminance and chromatic stimuli. *Journal of the Optical Society of America A*, 7, 128–140.

D'Zmura, M. (1991). Color in visual search. *Vision Research*, 31, 951–966.

D'Zmura, M., & Knoblauch, K. (1998). Spectral bandwidths for the detection of color. *Vision Research*, 38, 3117–3128.

D'Zmura, M., Lennie, P., & Krauskopf, J. (1987). Hue selectivity revealed by heterochromatic noise masking. *Investigative Ophthalmology & Visual Science*, 28(Suppl.), 92 (Abstract).

Dacey, D. M., & Lee, B. B. (1994). The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type. *Nature*, 367(Suppl.), 731–735 (Abstract).

Eskew, R. T., Jr., & Kortick, P. M. (1994). Hue equilibria compared with chromatic detection in 3D cone contrast space. *Investigative Ophthalmology & Visual Science*, 34(Suppl.), 1555 (Abstract).

Eskew, R. T., Jr., McLellan, J. S., & Giulianini, F. (1999). Chromatic detection and discrimination. In K. Gegenfurtner, & L. T. Sharpe, *Color vision: from genes to perception* (pp. 345–368). Cambridge: Cambridge University Press.

Eskew, R. T., Jr., Stromeyer, C. F., III, & Kronauer, R. E. (1992). The constancy of equiluminant red–green thresholds examined in two color spaces. *Advances in Color Vision Technical Digest (Optical Society of America, Washington DC)*, 4, 195–197.

Eskew, R. T., Jr., Stromeyer, C. F., III, & Kronauer, R. E. (1994). Temporal properties of the red-green chromatic mechanism. *Vision Research*, 34, 3127–3137.

Eskew, R. T., Jr., Stromeyer, C. F., III, Picotte, C. J., & Kronauer, R. E. (1991). Detection uncertainty and the facilitation of chromatic detection by luminance contours. *Journal of the Optical Society of America A*, 8, 394–403.

Gegenfurtner, K., & Kiper, D. (1992). Contrast detection in luminance and chromatic noise. *Journal of the Optical Society of America*, 9, 1880–1888.

Gegenfurtner, K. R., Kiper, D. C., & Levitt, J. B. (1997). Functional properties of neurons in macaque area V3. *Journal of Neurophysiology*, *77*, 1906–1923.

Giulianini, F. (1998). The yellow-blue detection mechanism as revealed by chromatic noise masking. Unpublished PhD thesis, Northeastern University, Boston.

Giulianini, F., & Eskew, R. T., Jr. (1998). Chromatic masking in the (ΔL/L,ΔM/M) plane of cone-contrast space reveals only two detection mechanisms. *Vision Research*, *38*, 3913–3926.

Graham, N. V. S. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. Huntington, NY: Krieger.

Ingling, C. R. (1977). The spectral sensitivity of the opponent-color channels. *Vision Research*, *17*, 1083–1089.

Kay, S. M. (1998). Detection theory. In *Fundamentals of statistical signal processing*, vol. II. Upper Saddle River, NJ: Prentice Hall.

Kiper, D. C., Fenstemaker, S. B., & Gegenfurtner, K. R. (1997). Chromatic properties of neurons in macaque area V2. *Visual Neuroscience*, *14*, 1061–1072.

Krauskopf, J., Williams, D. R., & Heeley, D. W. (1982). Cardinal directions of color space. *Vision Research*, *22*, 1123–1131.

Krauskopf, J., Williams, D. R., Mandler, M. B., & Brown, A. M. (1986). Higher order color mechanisms. *Vision Research*, *26*, 23–32.

Lennie, P., Krauskopf, J., & Sclar, G. (1990). Chromatic mechanisms in striate cortex of macaque. *Journal of Neuroscience*, *10*, 649–669.

Li, A., & Lennie, P. (1997). Mechanisms underlying segmentation of colored texture. *Vision Research*, *37*, 83–97.

McLellan, J. S., & Eskew, R. T., Jr. (2000). ON and OFF S-cone pathways have different long-wave cone inputs. *Vision Research*, *40*, 2449–2465.

McLellan, J. S., Goodman, J. B., & Eskew, R. T., Jr. (1994). Achromatic and chromatic detection of mixtures of blobs and isolated edges. *Investigative Ophthalmology & Visual Science*, *34*(Suppl.), 1370 (Abstract).

Mollon, J. D., Cavonius, C. R. (1987). The chromatic antagonisms of opponent process theory are not the same as those revealed in studies of detection and discrimination. In G. Verriest, *Colour vision deficiencies VIII* (pp. 473–483). Dordrecht: Junk.

Mullen, K. T., & Kulikowski, J. J. (1990). Wavelength discrimination at detection threshold. *Journal of the Optical Society of America A*, *7*, 733–742.

Pelli, D. (1990). The quantum efficiency of vision. In C. Blakemore, *Vision: coding and efficiency* (pp. 3–24). Cambridge: Cambridge University Press.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A*, *2*, 1508–1532.

Powell, I. (1981). Lenses for correcting chromatic aberration of the eye. *Applied Optics*, *20*, 4152–4155.

Pugh, E. N., Jr., & Mollon, J. D. (1979). A theory of the π1 and π3 color mechanisms of Stiles. *Vision Research*, *19*, 293–312.

Quick, R. F., Jr. (1974). A vector-magnitude model of contrast detection. *Kybernetik*, *16*, 65–67.

Sankeralli, M. J., & Mullen, K. T. (1996). Estimation of the L-, M-, and S-cone weights of the postreceptoral detection mechanisms. *Journal of the Optical Society of America A*, *13*, 906–915.

Sankeralli, M. J., & Mullen, K. T. (1997). Postreceptoral chromatic detection mechanisms revealed by noise masking in three-dimensional cone contrast space. *Journal of the Optical Society of America A*, *14*, 2633–2646.

Shinomori, K., Spillmann, L., & Werner, J. S. (1999). S-cone signals to temporal OFF-channels: asymmetrical connections to postreceptoral chromatic mechanisms. *Vision Research*, *39*, 39–49.

Stromeyer, C. F., III, Chaparro, A., Rodriguez, C., Chen, D., Hu, E., & Kronauer, R. E. (1998). Short-wave cone signal in the red-green detection mechanism. *Vision Research*, *38*, 813–826.

Stromeyer, C. F., III, Cole, G. R., & Kronauer, R. E. (1985). Second-site adaptation in the red-green chromatic pathways. *Vision Research*, *25*, 219–237.

Stromeyer, C. F., III, & Lee, J. (1988). Adaptational effects of short wave cone signals on red–green chromatic detection. *Vision Research*, *28*, 931–940.

Stromeyer, C. F., III, Thabet, R., Chaparro, A., & Kronauer, R. E. (1999). Spatial masking does not reveal mechanisms selective to combined luminance and red–green color. *Vision Research*, *39*, 2099–2112.

Thornton, J. E., & Pugh, E. N., Jr. (1983). Relationship of opponent-colours cancellation measures to cone-antagonist signals deduced from increment threshold data. In J. D. Mollon, & L. T. Sharpe, *Colour vision: physiology and psychophysics* (pp. 361–373). London: Academic Press.

Vos, J. J. (1978). Colorimetric and photometric properties of a 2° fundamental observer. *Color Research and Application*, *3*, 125–128.

Wandell, B. A. (1985). Color measurement and discrimination. *Journal of the Optical Society of America A*, *2*, 62–71.

Watanabe, A., Smith, V. C., Pokorny, J. (1997). Psychometric functions for chromatic discriminations. In C. R. Cavonius, *Colour vision deficiencies XIII* (pp. 369–376). Dordrecht: Kluwer Academic Publishers.

Watson, A. B. (1979). Probability summation over time. *Vision Research*, *19*, 515–522.

Watson, A. B., & Robson, J. G. (1981). Discrimination at threshold: labelled detectors in human vision. *Vision Research*, *21*, 1115–1122.

Webster, M. A., & Mollon, J. D. (1994). The influence of contrast adaptation on color appearance. *Vision Research*, *34*, 1993–2020.

Werner, J. S., & Wooten, B. R. (1979). Opponent chromatic mechanisms: relation to photopigments and hue naming. *Journal of the Optical Society of America*, *69*, 422–434.

Zaidi, Q., & Halevy, D. (1993). Visual mechanisms that signal the direction of color changes. *Vision Research*, *33*, 1037–1051.

Zaidi, Q., Shapiro, A., & Hood, D. (1992). The effect of adaptation on the differential sensitivity of the S-cone color system. *Vision Research*, *32*, 1297–1318.