

## COLOR DISCRIMINATION AT THRESHOLD: THE APPROACH THROUGH INCREMENT THRESHOLD SENSITIVITY

D. B. KIRK

Vision Research Laboratory, 5044 Kresge II, Box 56, The University of Michigan,  
Ann Arbor, MI 48109 U.S.A.

(Received 24 April 1981; in revised form 30 November 1981)

**Abstract**—A close correlation between Stiles's increment threshold ( $\pi$ -mechanism) approach and color discrimination at threshold was found: a sharp transition from detection of two lights via two  $\pi$ -mechanisms ( $\pi_1$  and  $\pi_4$ ) to detection via only one  $\pi$ -mechanism ( $\pi_1$ ) was paralleled by a sharp decline from excellent discrimination between those two lights (even with the rods bleached) to no discrimination. This parallel between detection and discrimination suggests that Stiles's approach can isolate single visual pathways with distinct perceptual correlates.

### INTRODUCTION

In his classic 1939 investigation, W. S. Stiles found a surprisingly systematic relationship between the wavelength and radiance of a large, steadily viewed background field, and the wavelength and threshold radiance of a  $1^\circ$ , 63 msec test flash. In plotting foveal log test threshold vs log field radiance curves (t.v.r. curves), Stiles (1939) found three distinct component branches, each branch having a fixed shape regardless of the test or field wavelength. Each branch obeyed the invariance properties that changes in the test wavelength caused shifts of each branch strictly parallel to the log test radiance axis, and similarly changes in the field wavelength caused shifts parallel to the log field radiance axis. Stiles (1953) later termed these branches the  $\pi_1$ ,  $\pi_4$  and  $\pi_5$  mechanisms. Although one additional branch was later found ( $\pi_2$ ; Stiles, 1953) in some observers' data, and three additional branches ( $\pi_3$ ,  $\pi_4$  and  $\pi_5$ ; Stiles, 1953, 1959) had to be admitted when high field radiances were employed, all seven branches had approximately the same shape and all obeyed the above invariance properties.

A  $\pi$ -mechanism is defined to be a branch in a t.v.r. curve obeying the above invariance properties, so that these mechanisms are strictly only elements in a mathematical analysis of increment threshold data. However, current evidence indicates that these mechanisms are closely related to many perceptual phenomena as well, e.g. metameric matching (Pugh and Sigel, 1978), brightness matching (Whittle, 1973), hue cancellation and chromatic adaptation (Larimer, 1981; Pugh and Larimer, 1980; Pugh and Mollon, 1978; Thornton, 1981), and color naming (Krauskopf, 1978). Perceptual correlates for these mechanisms have also been described. For example, Stiles (1949, p. 152) observed that "... in passing from the low to the high intensity range of t.v.i. [or t. v. r.] curves showing the transition from the "green" [ $\pi_4$ ] to the "blue" [ $\pi_1$ ] mechanism ... the appearance of the test

stimulus at just above the threshold changes. The apparent colour becomes purplish instead of blue or blue-green and there is a loss in sharpness of outline." Similarly, Whittle (1974, p. 599) observed that as a violet flash is increased in radiance and crosses the  $\pi_4$  or  $\pi_4$  threshold, "... the flash changes hue and acquires the crisp appearance characteristic of vision with "green" or "red" mechanisms."

One particularly simple and appealing basis for such correlates and relations would be for each  $\pi$ -mechanism to represent the threshold response of a single unidimensional channel or visual pathway. Such a relationship would yield the following application to color discrimination between threshold flashes (i.e. discrimination between two equally detectable stimuli differing only in their spectral composition). Signals equated for the *same* channel (i.e.  $\pi$ -mechanism) would not be discriminable, whereas signals equated for *different* channels might well be easily discriminable. Such a relationship is supported by the preliminary experiments of Stiles (1949, cf. below) which suggest that lights detected only via  $\pi_1$  are not discriminable, but that lights detected separately via  $\pi_1$  and  $\pi_4$  are discriminable. Similarly, Rollman and Nachmias (1972, cf. below) found that two lights, one which should be detected via  $\pi_4$  and the other via  $\pi_5$ , are discriminable.

A further test of this relationship can be made if, for a fixed field (i.e. steady background) wavelength, two test flash wavelengths can be chosen so that a change in the field radiance causes a fairly sharp shift from detection of each flash via *different*  $\pi$ -mechanisms to detection via the *same*  $\pi$ -mechanism. Assuming good discrimination between the flashes when detected via different  $\pi$ -mechanisms, the sharp transition in detection caused by the field radiance change should be matched by a similarly sharp decline in discrimination between the two test flashes.

Using appropriate test and field wavelengths, such a sharp transition of detection via  $\pi_4$  and  $\pi_1$  to detec-

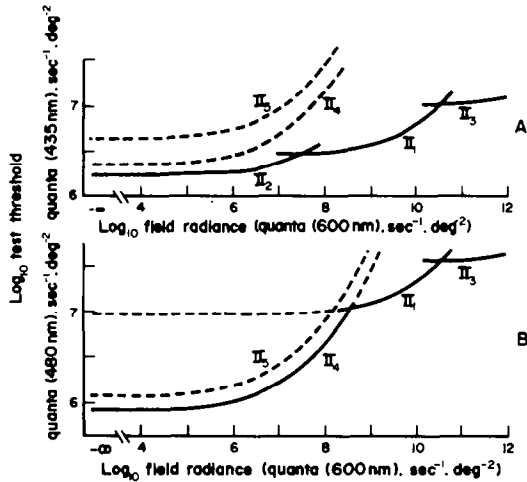


Fig. 1. Panel A gives predicted  $\pi$ -mechanism thresholds for a 435 nm, 200 msec,  $1^\circ$  foveal test flash upon 600 nm background fields. Predictions are based on the field sensitivities, Weber fractions, and standard template given in Wyzecki and Stiles (1967, Tables 7.4–7.6). Panel B gives similar predictions for a 480 nm test flash upon 600 nm background fields.

tion via  $\pi_1$  only can be obtained. For example, using 600 nm background fields, Stiles obtained approximately the lower envelope of the solid curves in Fig. 1a and 1b as the thresholds for 435 and 480 nm flashes, respectively. The dashed lines show the predicted positions of other  $\pi$ -mechanisms based on the invariance properties given earlier. Under such conditions, the 435 nm flash was detected via  $\pi_2$  for low field radiances and via  $\pi_1$  for higher field radiances. For the 480 nm flash, a sharp transition between the  $\pi_4$  and  $\pi_1$  mechanisms occurs, so that slightly above about  $8.5 \log \text{ quanta sec}^{-1} \text{ deg}^{-2}$  both the 480 and 435 nm flashes were detected via  $\pi_1$  only.

Thus by measuring the degree of discrimination between the 480 and 435 nm flashes at various radiances of the 600 nm field it is possible to determine if changes in discrimination (if any) are correlated with transitions between detection via different  $\pi$ -mechanisms. In particular, one can test if equally detectable flashes, one flash detected via  $\pi_4$  and the other via  $\pi_1$  or  $\pi_2$ , are discriminable. Similarly, one can test if equally detectable flashes detected only via  $\pi_1$  are discriminable. Such a test of discrimination between the 435 and 480 nm flashes detected via  $\pi_1$  should be very nearly the most sensitive possible test of such discrimination, since wavelengths appreciably longer than 480 nm cannot be used if  $\pi_1$  is to be isolated.

Discrimination between flashes detected via  $\pi_4$  and  $\pi_1$  or  $\pi_2$  was found to be excellent, whereas discrimination between flashes detected only via  $\pi_1$  was very poor or absent. The sharp t.v.r. curve transition between detection via  $\pi_4$  and  $\pi_1$  to detection via  $\pi_1$  alone was indeed found to be well-matched with a similarly sharp decline in discrimination.

## METHODS

### Apparatus

The 4-channel Maxwellian view apparatus used in this experiment is described in detail in Shevell (1978). Briefly, the light source was a 600 watt and later a 700 watt xenon lamp (Hanovia 981-C1 and 991-C1, respectively), which illuminated four Schoeffel GM 100 double monochromators. Only three of the double monochromators were used; two of them provided the 435 and 480 nm test lights and the third provided the 600 nm background field (M1, M2 and M3, respectively, from Shevell, 1978). Light from each of the monochromators was attenuated with balanced 2 log unit Wratten neutral density wedges; the 435 and 480 nm lights were set to threshold by the observer or investigator (depending on the method used for setting threshold) moving the wedges for the respective lights via torque transmitting and receiving synchros. After this attenuation, the 480 and 435 nm lights were then mixed and passed through a common shutter (at a focal point), the shutter controlled electronically so that 200 msec flashes of light could be presented at regular intervals. The 480 or 435 nm light could be extinguished by lowering (without detectable sound or vibration) a flap over the respective monochromator; the use of a common shutter, as well as a common  $3/4^\circ$  field stop, prevented the observers from discriminating the flashes with auditory, vibration, or external spatial cues. All lenses used were achromats, except for an achromatizing lens used during part of the experiment to correct for the axial chromatic aberrations of the average human eye (Wyzecki and Stiles, 1967, p. 212). The image formed at the pupil plane was less than 2 mm in diameter, so that no pupil dilation was necessary. The observer's head was held firmly fixed by a bite bar with a full dental impression, the bar being mounted in a three-dimensional positioning device. When the background field radiance was low, fixation was maintained by fixating between 4 small fixation lights (located at the  $3/4^\circ$  field stop) in a diamond pattern with each light  $2.5'$  from the center of fixation. For higher field radiances, the observer fixated on a small speck in the center of the  $5^\circ$  background field.

The bleaching source used was a 12 V tungsten coil filament, which was imaged onto an aperture stop, with the stop adjusted to produce a 2 mm image centered at the center of the observer's pupil. After passing through the aperture stop, the light was collimated and reflected into the background field channel.

### Calibration

The absolute luminance levels at 580 nm were measured with a MacBeth illuminometer as described in Shevell (1978). Using these measurements, the radiance levels for the stimuli actually used were then calculated using a United Detector Technology PIN-10DF photodiode/filter combination. The PIN-10DF is designed to have a flat spectral response

( $\pm 10\%$ ) from 450 to  $>900$  nm (for an equal energy spectrum), which was verified by calibrating the photodiode from 400 to 700 nm with a United Detector Technology UDT-500 photodiode, which in turn was calibrated by the manufacturer. The Wratten neutral density wedges were calibrated with the PIN-10DF for each wavelength of light they transmitted in the experiment, and for every density used, except for the wedges used in setting the 435 and 480 nm flash thresholds. Here the density was measured every 0.05 log units, linearly interpolating for density values not directly measured.

Photodiode readings were made at the end of each experimental session, and immediately after the absolute calibration session in which the MacBeth illuminometer was used, so the between session change in photodiode voltage, and thus the correct absolute radiance levels, could be calculated. Within session variability was monitored and found to be negligible.

#### *Procedure*

Two observers participated in this study; both observers used their right eye and had normal color vision as tested by the Farnsworth–Munsell 100 hue test and anomaloscope matching, and neither required corrective lenses. They were told about their performance during and after sessions requiring discrimination between test flashes, but were kept naive of the purpose and theoretical framework of the experiment. The first observer, M.B., was a 25-yr old male with extensive experience as a psychophysical observer, and the second observer, D.W., was a 20-yr old female who had not previously participated in a psychophysical experiment.

Observers were aligned in the apparatus by a combination of (1) determining if the tungsten filament image of the bleaching light fell on the observer's sclera when the observer looked to one side, (2) a rough determination of the Stiles-Crawford effect of the first kind, and (3) by centering the field in its "halo" of scattered light. No attempt was made to dark adapt the observer at the beginning of any session; either the observer's eye was bleached (see below), or the observer was initially adapted to the background field for at least 10 min, practice threshold measurements being made during this time.

The stimuli employed were circular and centrally fixated, consisting of a  $3/4^\circ$ , 200 msec, 435 or 480 nm test flash concentric with a steady  $5^\circ$  600 nm background field. Threshold vs radiance (t.v.r.) curves similar to those of Stiles (cf. Fig. 1a and 1b) were obtained for M.B. by first bleaching his eye for 5 min with a 5.65 log td tungsten "white" light subtending slightly more than  $10^\circ$  of visual angle (to eliminate rod contribution to threshold). Then following 5 min of dark adaptation, a full increment threshold curve was obtained, allowing about 1 min of adaptation time for the lowest 2 or 3 field radiances, increasing the adaptation time to 3–7 min for the higher field

radiances, always allowing threshold to stabilize before taking measurements. Threshold settings were always made using the method of adjustment, one 200 msec flash being presented every 1.5 sec.

Most of the discrimination experiments were done with M.B. as the observer; D.W.'s data were collected only to determine if the main features of M.B.'s data held for other observers, and no t.v.r. curves were obtained for D.W. For some of the discrimination sessions, M.B.'s (but not D.W.'s) eye was bleached in the same manner as when the t.v.r. curves were obtained (again to eliminate the rods), except that only one background field was used, introduced 2–3 min after the offset of the bleaching field. Measurements were taken from 6 to 15 min after offset of the bleaching field, a period in which rod threshold should be elevated more than about 1 log unit above their dark-adapted threshold.

The procedure for conducting the discrimination trials was the same regardless of whether or not the observer's eye had been bleached. The discrimination trials for a given field intensity were grouped into blocks, each block consisting of a practice session followed by the actual data gathering period in which thirty test flashes were presented, with 15 presentations of the threshold 480 nm test randomly intermixed with 15 presentations of the threshold 435 nm test. In addition, 5–10 blank trials were randomly intermixed in these 30 flash presentations in some sessions to obtain an estimate of the observer's false alarm rate (found to be between 5 and 10%). The practice session consisted of (1) 40 presentations of each threshold flash, alternating flash wavelengths in groups of 4, (2) 10 practice discrimination trials, and (3) about 15 presentations of each flash, as in (1). The purpose of the practice session was to enable the observer to learn the flash labels and apply them correctly, if possible, to each flash. During this practice session, the 480 nm flash was given the label A or B, the 435 nm flash was given the remaining label, and the observer was given the correct label (only A or B, and not any descriptive labels) as each light was presented. Labels were chosen randomly between each block of 30 trials, but were fixed within a block of trials.

Thirty discrimination trials were then given, the temporal sequence consisting of (1) the investigator raising the appropriate monochromator flap (if the trial was not a blank) and giving a verbal ready signal, (2) the observer actuating the shutter, thus opening it for 200 msec, (3) the observer pressing one of four buttons to signal if either (a) the flash was seen and was probably light A, or (b) the flash was seen and was probably light B, or (c) the flash was not seen, but was most probably A, or (d) the flash was not seen but was most probably light B, (4) the investigator both recording and repeating to the observer that response, and (5) similarly recording and informing the observer if the flash was actually light A, B, or a blank.

In order to start any given block of trials, it was necessary first to equate the thresholds for the two flashes so that both were seen about 50% of the time. At first, the method of double random staircases was used to equate threshold, but the method of adjustment was found to produce nearly identical results much more quickly. Occasionally, it was found that one or both of the 480 and 435 nm flash thresholds were definitely too high or too low (roughly more than 67% or less than 33% detected) when the block of discrimination trials was begun. If this was the case, or if the observer reported any brightness differences at all between the two flashes, the radiance of the flash(es) was changed in the appropriate direction by 0.05 to 0.15 log units, and the entire block of trials, including the practice periods and possible relabeling of stimuli, was started over. Threshold settings for a given field radiance (rounded to the nearest 0.1 log unit) always differed by less than 0.2 and 0.35 log units for within- and between-day measurements, respectively.

Discrimination for a block of 30 trials was deemed possibly significant if the Fisher's exact test value (Hays, 1973; Kendall and Stuart, 1961) was less than 0.25 for each of the three  $2 \times 2$  contingency tables formed from each block of trials. The dimensions within each table were flash type presented (435 or 480 nm) and flash type response, the tables differing only in that the first table contained only data in which the flash was reported detected, the second contained only data in which the flash was reported not detected, and the third contained all of the data regardless of whether or not the flash was reported detected.

If such a possibly significant amount of discrimination was found for a block of 30 trials, it was required that the detection frequencies for both flashes (the number of times each flash was detected out of 15 presentations of that flash) be  $\leq 9$ , so that both flashes were detected  $< 2/3$  of the time they were presented. It was also required that differences between the detection frequencies be  $\leq 3$ , so that discrimination could not be based on a brightness or detection frequency difference between the flashes. Although a significance level of 0.05 is usually the maximum significance level employed, the much higher level of 0.25 was employed here since the purpose was to exclude all data that are possibly significant due to one or both of the flashes being supra-threshold or not being equated for detectability.

To ensure that discrimination was not based on detection of differential chromatic aberration between the flash wavelengths, some discrimination trials were conducted with an achromatizing lens in place; no evidence for impairment of discrimination was found. To further ensure discrimination was based only on wavelength differences, discrimination trials were conducted with flashes of the same wavelength, and no evidence for discrimination was found.

## RESULTS

M.B.'s t.v.r. curves for  $\lambda = 480$  and 435,  $\mu = 600$  are plotted in Fig. 2, fitted by eye with Stiles's standard template (Wyszecki and Stiles, 1967, Table 7.5). The squares and triangles represent the thresholds for the 435 and 480 nm test flashes, respectively. The test and field sensitivities for  $\pi_1$ ,  $\pi_2$  and  $\pi_4$  fit Stiles's average values (c.f. Fig. 1) with a difference of less than the  $\pi_4$  and  $\pi_1$  test and field sensitivity standard deviations found by Stiles (1946, p. 431) (applying Stiles's  $\pi_1$  standard deviations to M.B.'s  $\pi_2$  as well as his  $\pi_1$ ). These agreements, as well as (1) the obvious flattening of M.B.'s t.v.r. curve obtained with a 480 nm test (due to the  $\pi_4$  to  $\pi_1$  transition), which Stiles (1946) found for all 20 observers, and (2) the strictly vertical shift of the probable  $\pi_1$  branch as the test was changed from 435 to 480 nm, together ensure that the branches are properly labeled. T.v.r. curves for  $\lambda = 480$  and 435 nm and  $\mu = 535$  and 638 nm were also obtained for M.B.: each  $\pi$ -mechanism was displaced (at least approximately) only parallel to the field radiance axis and agreed with Stiles's average values to within Stiles's (1946) 1 standard deviation values.

Note that, for Fig. 2, the  $\pi_4$  to  $\pi_1$  transition for detection of the 480 nm test occurs at 8.5 log quanta  $\text{sec}^{-1} \text{deg}^{-2}$ , as it did for Stiles's average curves in Fig. 1. Thus above 9.0 (understanding the units to be log quanta  $\text{sec}^{-1} \text{deg}^{-2}$  of the 600 nm field)  $\pi_1$  should be isolated from  $\pi_4$  by at least about 0.4 log units. Hence by selecting a field radiance between 9.0 and about 11.0, both the 480 and 435 nm test flashes should be detected only via  $\pi_1$ , so this range has been labeled the  $\pi_1, \pi_1$  range. Between 8.5 and 9.0, it is not clear if  $\pi_1$  will be isolated, so this ambiguous range has been labeled with a question mark. But for more moderate field radiances between 8.0 and 8.5, the 435 and 480 nm flashes should be detected via  $\pi_1$  and  $\pi_4$  respectively, so this range has been labeled the  $\pi_1, \pi_4$  range. For field radiances less than 8.0, the 435 and 480 nm lights should be detected via  $\pi_2$  and  $\pi_4$  respectively, so this range has been labeled the  $\pi_2, \pi_4$  range.

To increase the sensitivities of the statistical tests employed here, M.B.'s discrimination data were pooled over  $1/2$  log unit bins (discrimination data for

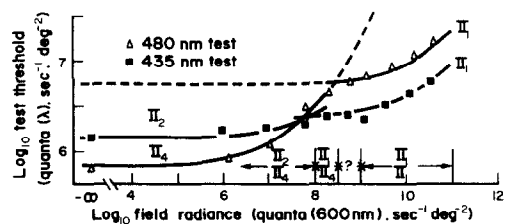


Fig. 2. T.v.r. curves for observer M.B. for a 435 (■) and a 480 (△) nm foveal test flash (200 msec,  $3/4^\circ$ ) upon 600 nm background fields. The templates are Stiles's standard template, as for Fig. 1. The respective  $\pi$ -mechanisms which the 435 and 480 nm flashes should stimulate are indicated above the abscissa.

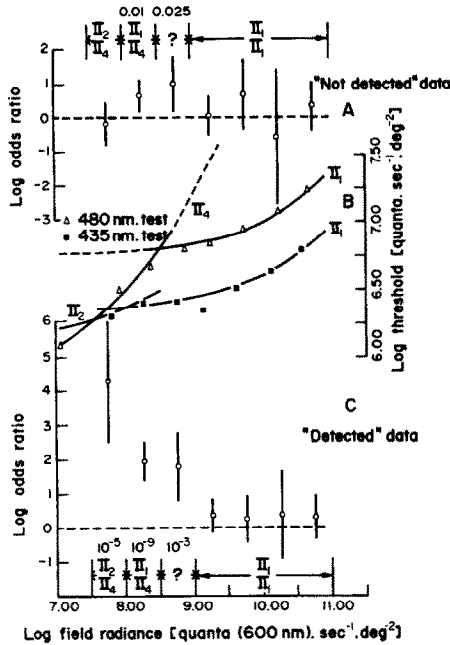


Fig. 3. Panels A and C give the log odds ratios, 95% confidence intervals, and significance levels obtained (for departures from 0) for those log odds ratios for each 1/2 log unit bin between 7.5 and 11.0 log quanta  $\text{sec}^{-1} \text{deg}^{-2}$ . Panels A and C are based only on those data in which observer M.B. reported not detecting the flash (Panel A) and reported detecting the flash (Panel C). Panel B has been taken from Fig. 2 for purposes of comparison, as has the  $\pi$ -mechanism notation for each bin. Directly above that notation are the obtained significance levels for each bin, except where  $P > 0.05$ . The dashed line at 0 gives the log odds ratio expected if no discrimination is possible between the flashes.

the  $\pi_2$ ,  $\pi_4$  (7.5–8.0),  $\pi_1$ ,  $\pi_4$  and 9.0–9.5 ranges were clustered at their respective range midpoints; data for all other bins were distributed roughly equally about their entire 1/2 log unit ranges). To test for significant discrimination, (natural) log odds ratios were computed for each bin. This odds ratio, call it  $\hat{a}$ , is simply the ratio of 435 to 480 nm responses given that the flash was actually 435 nm divided by the ratio of the 435 to 480 nm responses given that the flash was actually 480 nm. Thus if  $x(i, j)$  = the frequency for the type  $i$  response to the type  $j$  flash presentation (with type 1 = 435 nm and type 2 = 480 nm), then

$$\log \hat{a} = \log \left\{ \frac{x(1, 1)/x(2, 1)}{x(1, 2)/x(2, 2)} \right\} = \log \left\{ \frac{x(1, 1)x(2, 2)}{x(1, 2)x(2, 1)} \right\}$$

Two such log odds ratios were computed for each bin, one for the data in which M.B. reported detecting the flash (the "detected" data), and one for the data in which M.B. reported *not* detecting the flash (the "not detected" data). Pooling over 1/2 log unit bins allowed the sample sizes to be sufficiently large so that the sampling distribution of  $\log \hat{a}$  was normally distributed with mean  $\log a$ , the population log odds ratio of which  $\log \hat{a}$  is the maximum likelihood estimator (Everitt, 1977; Feinburg, 1977; Gart, 1971). The

sampling variance is then approximated by

$$s^2 = 1/x(1, 1) + 1/x(1, 2) + 1/x(2, 1) + 1/x(2, 2).$$

To test for independence (no discrimination),  $\log(\hat{a})/s$  can be compared to a normal (0, 1) distribution, as can  $(\log \hat{a}_1 - \log \hat{a}_2) / \sqrt{(s_1^2 + s_2^2)}$  to test for significant differences between log odds ratios  $\log \hat{a}_1$  and  $\log \hat{a}_2$ . Since the alternative to the null hypothesis of no discrimination is one of discrimination or positive  $\log \hat{a}$  values, the test for no discrimination is one-tailed. Since differences between different log odds ratios may be positive or negative, a test for such differences should be two-tailed.

Within either the  $\pi_1$ ,  $\pi_4$  or ambiguous ranges, no significant differences ( $P > 0.05$ ) were found between the log odds ratios when M.B.'s eye had or had not been bleached: for the "detected" data in these ranges, the log odds ratios were nearly equal, differing by less than 0.3. The data were thus pooled over the bleached and non-bleached conditions within each of these ranges. Such pooling was not necessary in the other ranges: all of the data in the  $\pi_2$ ,  $\pi_4$  range were obtained in the bleached condition, and the rods should be saturated in the  $\pi_1$ ,  $\pi_1$  range (at 600 nm, 9.0 log quanta  $\text{sec}^{-1} \text{deg}^{-2} = 3.8 \text{ log scot. td}$ ; Aguilar and Stiles, 1954).

In the lower panel of Fig. 3 are plotted the log odds ratios for M.B.'s "detected" data, along with 95% confidence intervals. The numbers directly below that plot give the obtained one-tailed significance level of departures from the null hypothesis of random responding (no discrimination). The middle panel in Fig. 3 gives t.v.r. curves from Fig. 2 for comparison. Discrimination is quite clearly excellent in the ( $\pi_4$ ,  $\pi_2$ ), ( $\pi_4$ ,  $\pi_1$ ) and ambiguous ranges, particularly in the  $\pi_4$ ,  $\pi_1$  range where there is less than a 1 in 1 billion chance that the null hypothesis of no discrimination is correct. But the data in the  $\pi_1$ ,  $\pi_1$  range indicate that discrimination is poor or absent here; all of the confidence intervals, even though fairly narrow, include 0, the value expected if there is no discrimination. Similarly, for the  $\pi_1$ ,  $\pi_1$  range there are no significant ( $P > 0.05$ ) departures from the null hypothesis of no discrimination.

Again using a two-tailed test for differences between the log odds ratio (as was used to test for differences in bleaching conditions), no significant differences ( $P > 0.05$ ) were found between the four  $\pi_1$ ,  $\pi_1$  1/2 log unit bins. The  $\pi_2$ ,  $\pi_4$  log odds ratio, however, is significantly different at the  $P = 0.025$  level from all of the other 6 log odds ratios. The ratios for the  $\pi_1$ ,  $\pi_4$  and ambiguous ranges are not significantly different from one another. However, both are significantly different (at the 0.0001 and 0.025 levels, respectively) from the ratio for the 9.0–9.5 range, indicating a rather large change in discrimination for a relatively small change in field radiance of about 1 and 0.5 log units respectively.

The log odds ratios and the significance levels obtained when M.B. reported not detecting the flash

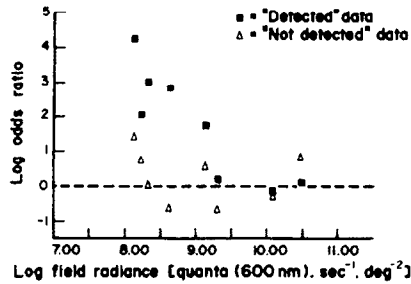


Fig. 4. Log odds ratios for observer D.W. for each block of 30 trials, obtained as for observer M.B. The squares and triangles represent those log odds ratios when D.W. reported detecting the flash (■) and not detecting the flash (△) respectively. The dashed line at 0 gives the log odds ratio expected if no discrimination is possible between the flashes.

are given in the upper panel of Fig. 3. For that "not detected" data, discrimination is significant in the  $\pi_4$ ,  $\pi_1$  and ambiguous ranges at the  $P = 0.01$  and  $0.025$  levels, respectively. None of the 7 log odds ratios for that "not detected" data are significantly different from each other at the  $P = 0.05$  level.

D.W.'s data are given in Fig. 4. Each point is the log odds ratio for an individual block of 30 trials, the squares representing those ratios for the "detected" data and the triangles representing those ratios for the "not detected" data. Each log odds ratio is based on response frequencies which are too small (zero entries being replaced with 0.5) to apply the normal approximations used to obtain the confidence intervals and significance levels for M.B.'s data. However, such approximations can be applied if the data are pooled between 8.2 and 8.6 and between 9.1 and 10.5 log quanta  $\text{sec}^{-1} \text{deg}^{-2}$ . For the first range of 8.2–8.6 (call it the  $\pi_1$ ,  $\pi_4$  range), the 480 nm flash should be detected via  $\pi_4$  and the 435 nm flash via  $\pi_1$  for either MB or Stiles's average observer. For the second range of 9.1–10.5 (call it the  $\pi_1$ ,  $\pi_1$  range) both flashes should be detected via  $\pi_1$ . The 95% confidence intervals and significance levels for the log odds ratios for those ranges are given in Table 1. The two ratios for the "not detected" data are not significantly different ( $P > 0.05$ ), but the two ratios for the "detected" data are significantly different at the  $P = 0.005$  level. This reflects the change from excellent discrimination ( $P < 10^{-6}$ ) in the  $\pi_1$ ,  $\pi_4$  range to the very poor or absent discrimination ( $P > 0.05$ ) in the  $\pi_1$ ,  $\pi_1$  range, quite similar to the transition found for observer M.B.

## DISCUSSION

The major finding of the present experiment is the close correlation between discrimination and the transition from detection via different  $\pi$ -mechanisms ( $\pi_1$  and  $\pi_4$ ) to detection via only one  $\pi$ -mechanism ( $\pi_1$ ). This correlation quite probably results from a transition from detection of the two test flashes via several classes of cones to detection via only one class. Assuming that each cone class is univariant (Naka and Rushton, 1966), i.e. assuming each class can signal only the rate quanta are absorbed and not the stimulus spectral composition, discrimination between two equally detectable lights differing only in their spectral composition implies differential signals between at least two cone classes. In particular, the excellent discrimination found between  $\pi_4$  and  $\pi_1$  (and  $\pi_4$  and  $\pi_2$ ) implies such differential cone inputs to those  $\pi$ -mechanisms.

Now, of course, the (contrapositive of the) converse of the above statement does not hold: zero discrimination between two lights does not imply detection via a single cone class. But in the case of discrimination between two lights detected only via  $\pi_1$ , a likely physiological mechanism for the above  $\pi_1$ -detection results is indeed detection via a single cone class—the short-wavelength-sensitive cones. In addition to the violet or purple hue sensation and diffuse temporal and spatial qualities of lights detected via  $\pi_1$  (Brindley, 1954; Brindley, Du Croz, and Rushton, 1966; Green 1968, 1969), such a strictly short-wavelength-sensitive cone input to  $\pi_1$  is strongly supported by (1) the sharp decline (roughly 1 log unit) of  $\pi_1$  test (and field) sensitivity from 435 to 500 nm (Stiles 1939, 1953; Pugh, 1976) whereas the middle- and long-wavelength sensitive cones should be increasing in sensitivity in this region, (2) the low field sensitivity of  $\pi_1$  (e.g.  $\pi_1$  threshold is elevated less than 2/3 log unit by a 570 nm field which should bleach about 1/2 of the middle- and long-wavelength-sensitive cone pigments), (3) the finding of Pugh and Sigel (1978) that the field sensitivity ( $\mu \leq 533$  nm) of  $\pi_3$  (and hence  $\pi_1$  test sensitivity) is well-fit by a linear combination of the Stiles and Burch (1955) small field color matching functions;  $\pi_3$  field and  $\pi_1$  test sensitivities have the same shape insofar as the latter can be measured (Pugh, 1976; Stiles 1939, 1953), and (4) the finding of Ikeda, Uetseki and Stiles (1970) that 420 and 470 nm flashes detected via  $\pi_1$  exhibit energy summation when the energy units for each flash are

Table 1. Log odds ratios and 95% confidence intervals (observer D.W.)

| Field radiance range<br>(log quanta<br>(600 nm)<br>$\text{sec}^{-1} \text{deg}^{-2}$ ) | "Detected" data |                         | "Not detected" data |                         |
|--|-----------------|-------------------------|---------------------|-------------------------|
|  | Log odds ratio  | 95% confidence interval | Log odds ratio      | 95% confidence interval |
| 8.2–8.6  | 2.8†            | $2.8 \pm 1.1$           | -0.20*              | $-0.20 \pm 1.0$         |
| 9.1–10.5   | 0.21*           | $0.21 \pm 1.3$          | 0.24*               | $0.24 \pm 1.2$          |

\*Denotes  $P > 0.05$ ; †denotes  $P < 10^{-6}$ .

equated at threshold: Ikeda *et al.* (1970), Boynton *et al.* (1964), Guth (1965, 1967), Stiles (1967) and Krauskopf (1974) have consistently found a rather complex inhibitory and excitatory relation between cone classes, so that any test additivity between cone classes would be exceptional.

This strictly short-wavelength-sensitive cone input to  $\pi_1$  is not contradicted by the finding of Pugh (1976) that  $\pi_1$  does not obey the property of field additivity. Those nonadditivities involved *fields* of wavelength  $\geq 550$  nm, but  $\pi_1$  test sensitivity can only be measured between roughly 400 and 500 nm (Stiles, 1953), so those nonadditivities do not enter into the argument here. Indeed, Pugh (1976) showed that  $\pi_1$  is field additive for fields of wavelength  $\leq 500$  nm, a finding consistent with the hypothesis that both the  $\pi_1$  field and test sensitivities for wavelengths  $\leq 500$  nm are short-wavelength-sensitive cone action spectra.

#### Related work

Stiles (1949, pp. 150–152) proposed a similar experiment to the present one, using 480 and 435 nm stimuli placed side by side and flashed simultaneously on a 600 nm background field. Assuming both flashes are detected via  $\pi_1$  at threshold, Stiles proposed that those flashes would be indiscriminable at threshold but would become discriminable if the radiance of both were increased to the point where the 480 nm flash is detected via  $\pi_4$  as well as  $\pi_1$  (cf. Whittle's (1974, p. 599) observations above). Stiles (1949, p. 151) found that "preliminary experiments on these lines indicate that the two stimuli do in fact appear similar over a certain range above the threshold, and begin to differ in apparent colour at about the expected intensity." The present study is, in part, a test of Stiles's proposition that the two lights detected via  $\pi_1$  are, indeed, not discriminable, a proposition that appears correct. Stiles's finding supports the close correlation found between  $\pi$ -mechanisms and color discrimination found in this study.

Rollman and Nachimas (1972) employed an experimental paradigm very similar to the present one of presenting one of two equally detectable flashes (or a blank) and having the observer respond (1) whether or not the flash was detected and (2) if the flash was color A or B. The 43 msec flashes used were either green (Wratten 65) or red (Wratten 29) and 10' in diameter. No background field was used (except for one observer) so that the green and red flashes should have been detected via  $\pi_4$  and  $\pi_5$ , respectively. Thus since discrimination was quite good, equally detectable flashes, one detected via  $\pi_4$  and the other via  $\pi_5$ , should be discriminable.

Rollman and Nachimas (1972) also found (in agreement with the present experiment) that, if the flashes are discriminable when they are reported detected, some discrimination is often evident even when the observer reports *not* detecting those flashes. Unless one supposes that the observer gives some "no detect" responses while in a detect state, this is evidence

against a low (and high) threshold, two-state theory of detection.

*Acknowledgements*—This investigation has benefited greatly from the support and criticisms of Dr Mathew Alpern and Dr David H. Krantz. Dr Krantz provided the original concept for this investigation, which was generously supported by his grant NSF GB-36642X, as well as NEI training grant EY07022 for vision research to the University of Michigan.

#### REFERENCES

- Aguilar M. and Stiles W. (1954) Saturation of the rod mechanism of the retina at high levels of stimulation. *Optica Acta* **1**, 59–65.
- Boynton R. M., Ikeda M. and Stiles W. S. (1964) Interactions among chromatic mechanisms as inferred from positive and negative increment thresholds. *Vision Res.* **4**, 87–117.
- Brindley G. S. (1954) The summation areas of human color-receptive mechanisms at increment threshold. *J. Physiol.* **124**, 400–408.
- Brindley G. S., Du Croz J. and Rushton W. A. H. (1966) The flicker fusion of the blue-sensitive mechanism of color vision. *J. Physiol.* **183**, 497–500.
- Everitt B. S. (1977) *The Analysis of Contingency Tables*. Wiley, New York.
- Feinburg S. E. (1977) *The Analysis of Cross-Classified Categorical Data*. MIT, Cambridge.
- Gart J. J. (1971) The comparison of proportions: a review of significance tests, confidence intervals, and adjustments for stratification. *Rev. Int. Stat. Inst.* **39**, 148–169.
- Green D. G. (1968) The contrast sensitivity of the colour mechanisms of the human eye. *J. Physiol.* **196**, 415–429.
- Green D. G. (1969) Sinusoidal flicker characteristics of the color-sensitive mechanisms of the eye. *Vision Res.* **9**, 591–601.
- Guth S. L. (1965) Luminance addition: general considerations and some results at foveal threshold. *Vision Res.* **7**, 319–328.
- Guth S. L. (1967) Nonadditivity and inhibition among chromatic luminances at threshold. *Vision Res.* **7**, 319–328.
- Hays W. L. (1973) *Statistics for Psychologists* (2nd edn). Holt, Rinehart & Winston, New York.
- Ikeda M., Uetsuki T. and Stiles W. S. (1970) Interrelations among Stiles  $\pi$ -mechanisms. *J. opt. Soc. Am.* **60**, 406–415.
- Kendall M. G. and Stuart A. (1961) *The Advanced Theory of Statistics*. Griffin, London.
- Krauskopf J. (1974) Interaction of chromatic mechanisms in detection. *Mod. Probl. Ophthalm.* **13**, 92–97.
- Krauskopf J. (1978) On identifying detectors. In *Visual Psychophysics and Physiology* (Edited by Armington J. C., Krauskopf J. and Wooten B. R.), pp. 283–295. Academic Press, New York.
- Larimer J. (1981) Red/green opponent colors equilibria measured on chromatic adapting fields: evidence for gain changes and restoring forces. *Vision Res.* **21**, 501–512.
- Naka K. I. and Rushton W. A. H. (1966) S-potentials from colour units in retina of fish (Cyprinidae). *J. Physiol.* **185**, 535–555.
- Pugh E. N. Jr (1976) The nature of the  $\pi_1$  mechanism of W. S. Stiles. *J. Physiol.* **257**, 713–747.
- Pugh E. and Larimer J. (1980) Test of the identity of the site of blue/yellow hue cancellation and the site of chromatic antagonism in the  $\pi_1$  pathway. *Vision Res.* **20**, 779–788.

- Pugh E. and Mollon J. (1979) A theory of the  $\pi_1$  and  $\pi_3$  color mechanisms of Stiles. *Vision Res.* **19**, 293–312.
- Pugh E. and Sigel C. (1978) Evaluation of the candidacy of the  $\pi$ -mechanisms of Stiles for color-matching fundamentals. *Vision Res.* **18**, 317–330.
- Rollman G. B. and Nachmias J. (1972) Simultaneous detection and recognition of chromatic flashes. *Percept. Psychophys.* **12**, 309–314.
- Shevell S. K. (1978) The dual role of chromatic backgrounds in color perception. *Vision Res.* **18**, 1649–1661.
- Sigel C. and Pugh E. (1980) Stiles's  $\pi_5$  color mechanism: Tests of field displacement and field additivity properties. *J. opt. Soc. Am.* **70**, 71–81.
- Stiles W. S. (1939) The directional sensitivity of the retina and the spectral sensitivity of the rods and cones. *Proc. R. Soc. B* **127**, 64–105.
- Stiles W. S. (1946) Separation of the 'blue' and 'green' mechanisms of foveal vision by measurements of increment thresholds. *Proc. R. Soc. B* **133**, 418–434.
- Stiles W. S. (1949) Increment thresholds and the mechanisms of color vision. *Documenta ophthalmol.* **3**, 138–163.
- Stiles W. S. (1953) Further studies of visual mechanisms by the two-color increment threshold technique. *Coloq. Probl. opt. Vis., Madrid* **1**, 65–103.
- Stiles W. S. (1959) Color vision: the approach through increment threshold sensitivity. *Proc. natn Acad. Sci. U.S.A.* **45**, 100–114.
- Stiles W. S. (1967) Mechanism concepts in colour theory. Third Newton Lecture *Proc. natn Acad. Sci. U.S.A.* **45**, 106–123.
- Stiles W. S. and Burch J. M. (1955) N.P.L.'s investigation of colour-matching: interim report. *Optica Acta* **4**, 168–181.
- Thornton J. E. (1981) Relating chromatic antagonism in  $\pi_5$  to red/green hue cancellation. Ph.D. thesis. University of Michigan. University Microfilms, Ann Arbor.
- Wandell B. and Pugh E. (1980) A field-additive pathway detects brief-duration, long-wavelength incremental flashes. *Vision Res.* **20**, 613–624.
- Whittle P. (1974) Intensity discrimination between flashes which do not differ in brightness: some new measurements on the "blue" cones. *Vision Res.* **14**, 599–601.
- Wyszecki G. and Stiles W. S. (1967) *Color Science*. Wiley, New York.