Vision Research Vol. 19, pp. 293–312
© Pergamon Press Ltd 1979. Printed in Great Britain
0042-6989/79/0215-0293$02.00/0

# A THEORY OF THE $\Pi_1$ AND $\Pi_3$ COLOR MECHANISMS OF STILES

E. N. PUGH Jr and J. D. MOLLON*
Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, U.S.A. and
Psychological Laboratory, University of Cambridge, Cambridge, England

(Received 6 April 1978; in revised form 20 August 1978)

**Abstract**—A unified account of the $\Pi_1$ and $\Pi_3$ color mechanisms of Stiles is proposed. We hypothesize that under the conditions that isolate these branches of the two-color threshold a signal originating with the photon absorptions of the short-wavelength cones passes through two sites of attenuation. The first site's gain is controlled by the short-wavelength cones alone; the second site's gain is determined by a net "blue/yellow" opponent signal. The two sites are distinguished by spectral sensitivity, absolute field sensitivity and dynamic properties. The theory, formalized in four equations, provides a good account of these features of the increment thresholds: (1) super-additivity of the effects of $\Pi_1$-equated long- and short-wavelength fields; (2) upward deviation of the $\Pi_1$ threshold from the Weber line on bright blue backgrounds; (3) cancellative subadditivity of fields whose mixture is in approximate blue/yellow equilibrium; (4) the "limited conditioning effect" of long-wavelength fields; (5) transient "tritanopia".

## INTRODUCTION

The purpose of this paper is to offer a unified account of a number of striking effects that have been attributed to the short-wavelength sensitive or "blue" cones of human color vision. Most of the phenomena to be discussed, including results of our own investigations, have been observed and measured in variants of the two-color increment threshold experiments of W. S. Stiles (1939, 1949a, 1953, 1959, 1978). The essence of our hypothesis is that under the specific set of conditions that isolate two of the short-wavelength sensitive branches of the two-color threshold curves ($\Pi_1$ and $\Pi_3$), visual signals initiated by photons absorbed by the short-wavelength sensitive cones must pass successively through two distinct "sites" where they may be attenuated. Attenuation at the first site is determined exclusively by photons absorbed in short-wavelength cones themselves; attenuation at the second site is determined by "opponent" or antagonistically coded signals from the short-wavelength cones and the other cone classes. The hypothesized sites are defined and distinguished operationally by both steady-state (e.g. spectral) and kinetic properties. We conclude that this single pathway (comprising the short-wave cones and the two sites) accounts for the existence and behavior of the $\Pi_1$ and $\Pi_3$ branches.

The organization of the paper is as follows: Firstly, we present a statement of our definitions and assumptions. Secondly, we discuss the critical experiments from which the theory was derived. Thirdly, we give a formal statement of the theory, and show how it can account for six distinct features of the increment threshold data. Finally, we discuss the relationship of our theoretical account to current physiological knowledge.

## I. ASSUMPTIONS: INTERPRETATION OF THE LAWS OF COLOR VISION

### *The laws of Trichromacy and Additivity of color matches: three cone classes initiate color vision*

The absorption of light quanta by three distinct visual pigments in three classes of cone photoreceptor is generally held to be the initial event in human color vision. Yet, in a fundamental sense, the notion *cones* refers primarily not to anatomical objects, but to entities inferred to explain the laws of color vision, and in particular to explain Grassmann's Laws of Color Mixture (Brindley, 1970; Krantz, 1975a). Thus, the Invariance of color matches under neutral attenuation, and the Additivity of matches[1] are thought to result from the "univariance" and spectral additivity properties of photopigment absorptions in several homogeneous classes of photoreceptor. And indeed, the Trichromacy of color matches remains the most compelling evidence that there are exactly three cone pigments. In short, while we assume three cone classes as the initial stage of color vision, we emphasize the

[1] Complete statements of Grassmann's Laws of Metameric Matching may be found in Wyszecki and Stiles (1967, pp. 228–234), and in Krantz (1975a). Since we refer often in this paper to the Laws of Invariance and Additivity, we present here a succinct statement. Let

| | |
|---|---|
| $A, B, C, \ldots$ | be arbitrary spectral distributions; |
| $*$ | represent *neutral attenuation*, i.e. for a scalar $t > 0$, $t * A$ is the spectrum obtained by multiplying all components by the factor $t$; |
| $\oplus$ | represent *physical superposition*, i.e. $A \oplus B$ is the spectrum obtained by point-by-point addition of components; |
| $\equiv$ | represent *metameric match*, i.e. $A \equiv B$ means that $A$ and $B$ are indistinguishable to photopic vision. |

Then we have the laws:

*Invariance*: $A \equiv B$ implies $(\forall t > 0)\ t * A \equiv t * B$;
*Additivity*: $A \equiv B$ if and only if $(\forall C)$ $A \oplus C \equiv B \oplus C$.

empirical laws that are the foundation for this belief, for these laws provide the tools for testing the hypothesis that a single cone class determines a given visual mechanism.

*Testing the hypothesis that a color mechanism is determined by a single class of cone*

A color mechanism (Stiles, 1967; Krantz, 1975a) in general terms may be defined as a real-valued function (operating on the space of spectral irradiance distributions) that could serve photopic discriminability: if two spectral irradiance distributions result in different function values, the lights in question are not metameric. Two necessary conditions that a color mechanism must satisfy in order to be a candidate for one of the three cones of normal color vision are invariance and additivity.[2] Failure of a mechanism to exhibit these properties is thus sufficient cause to reject the hypothesis that it is a cone. In addition to the two tests of cone candidacy deducible from the Grassmann Invariance and Additivity properties, a third test of the hypothesis that a mechanism is a cone may be deduced from Additivity (Brindley, 1957; Krantz, 1975a): an additive mechanism must have a spectral sensitivity curve that is a linear combination of any color matching functions. A fourth test of the hypothesis follows from the assumption that the Grassmann Invariance and Additivity properties are determined at the level of photon absorptions: a mechanism that has a spectral sensitivity function that is severely different from any that can be generated by reasonable photopigments can be rejected from among the ranks of cone candidates.

*Stiles's $\Pi$-mechanisms: branches that obey displacement Laws*

Stiles's experiments have revealed multiple (5–7) branches in the two-color increment-threshold curves. In Stiles's operational terminology, a "$\Pi$-mechanism" is identified with a unique branch of a two-color increment threshold curve, a branch that obeys, to a reasonable approximation, Stiles's invariance or Displacement Laws (Stiles, 1946b, p. 46).[3] The simplest hypothesis, that each branch or $\Pi$-mechanism represents adaptation controlled by a single class of cone, is contradicted by the very existence of 5–7 branches with distinct field spectral sensitivities. Indeed, it is precisely the existence of three distinct branches whose peak spectral sensitivity lies in the short-wavelengths that forms the principal topic of this paper. However, the fact that not all seven $\Pi$-mechanisms can be identified with cones does not reveal which, if any, are viable candidates, and so direct experimental tests of the hypothesis are required in each case. Our theory's purpose is to explain the facts known to date about the $\Pi_1$ and $\Pi_3$ branches of the two-color threshold in terms of a few elemental constructs. These constructs provide formal explanations of the phenomena, and are not at this time meant to be assigned particular physiological interpretations. To assist the reader with our use of terms and symbols, we have supplied an appendix (Appendix I); the symbolic notation follows closely that used by Stiles.

## II. REJECTION OF THE HYPOTHESIS THAT $\Pi_1$ ADAPTATION IS DETERMINED BY ONLY ONE CONE CLASS

*The $\Pi_1$ and $\Pi_3$ branches of the two-color threshold: the defining results*

Figure 1 shows a set of two-color increment threshold curves obtained from one observer over a period of about six months in one of our laboratories. The observer's task was to detect a 1° diameter, foveally presented test flash of 200 msec duration and wavelength $\lambda = 425$ nm in the presence of a series of increasingly intense 10° backgrounds of wavelength $\mu = 590$ nm. The three "branches" observed are labelled according to Stiles's nomenclature.

Each branch has a unique test and field sensitivity: the test sensitivity, $\pi_{j\lambda}$, is the reciprocal threshold intensity for a branch at zero field intensity; the field sensitivity, $\Pi_{j\mu}$, is the reciprocal field intensity that causes a 1 log unit increase in threshold. The reciprocal field sensitivities of the branches are indicated by arrows drawn to the abscissa, and are near those of Stiles's average observer (Fig. 2) for these branches. The lowest branch, $\Pi_2$, occupies the smallest threshold range and is therefore the most difficult to isolate and study. Indeed, it is not found in all observers (Stiles, 1953), and may be somewhat labile within a given observer. We shall not discuss $\Pi_2$ further in this paper. In brief, then, we have found the $\Pi_1$ and $\Pi_3$ branches of the two-color threshold readily isolable, and the quantitative features described by Stiles quite general and reproducible (see also Fig. 5). We turn now to the testing of hypotheses about their origin.

*Testing the hypothesis that the adaptation of the $\Pi_1$ branch is determined by a single cone class*

*Spectral sensitivity and color matches.* The $\Pi_1$ field action spectrum (Fig. 2) has a secondary mode at long-wavelengths. In a population of 18 color-normal observers studied by Stiles (1946a) the range of the observed field sensitivity $\Pi_{1\mu}$ in the neighborhood of this secondary mode is unduly large, being twice the

---

[2] A *color mechanism* or *color code* is a real-valued function $f$ defined on the space of metameric lights: i.e.

$$f(A) \neq f(B) \text{ implies (not } A \equiv B).$$

Then for a mechanism $f$,

$f$ is *invariant* if and only if
$[f(A) = f(B) \text{ implies } (\forall t > 0)\, f(t * A) = f(t * B)]$;

$f$ is *additive* if and only if
$[f(A) = f(B) \text{ if and only if } f(A \oplus C) = f(B \oplus C)$.

Note that the additive property does *not* imply $f(A \oplus C) = f(A) + f(C)$. Krantz (1975a) suggests the term "Grassmann" be used instead of "additive"; Brindley (1957) used the term "substitutable" for a mechanism satisfying this property. Note that these properties of a color mechanism are distinct from the related properties of matches (Footnote 1). For details, see Krantz (1975a).

[3] The simplest statement of the Field Displacement Law (which is the only Displacement law explicitly referred to in the paper) is this: threshold elevation of the $j^{th}$($j = 0, 1, \ldots, 7$) branch is an *invariant color mechanism.* (See footnote 2). The term "displacement" arises because increment threshold curves are plotted on logarithmic axes; scalar multiplication is converted to displacement. (See Stiles, 1939, pp. 86–87.)

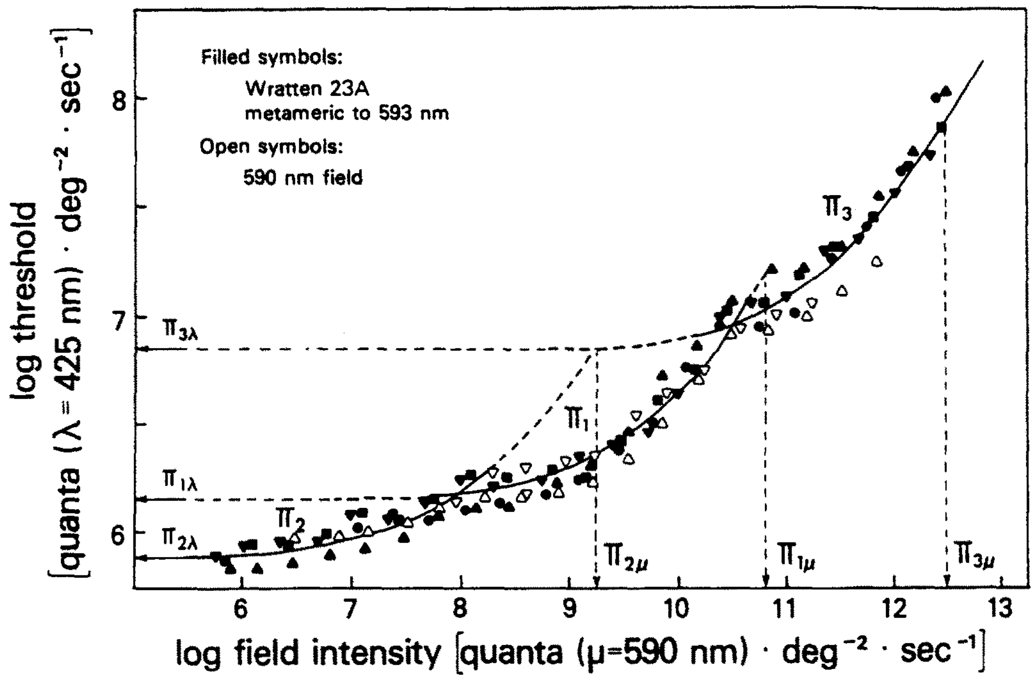


Fig. 1. A series of two-color foveal increment threshold curves for Observer EP for a 200 msec test ($\lambda$ = 425 nm and $\mu \simeq$ 590 nm), revealing the three short-wavelength-sensitive branches, $\Pi_2$, $\Pi_1$ and $\Pi_3$. In order to achieve intensities high enough to reveal the uppermost branch in some of the experiments a broad-band (high pass) Wratten 23A filter was used to color the light from the 75 W xenon source; this spectrum was *metameric* at 1000 td to a monochromatic light of 593 nm. The uppermost branch is less sensitive to the narrow-band 590 nm light, presumably because the receptor class that determines this branch does not participate in the match at 1000 td. The test sensitivities ($\pi_{j\lambda}$) and field sensitivities ($\Pi_{j\mu}$) of the branches are indicated by the arrows. Different symbols represent experiments carried out on different days over a period of about six months.

range of the test sensitivity $\pi_{1\lambda}$ at $\lambda$ = 480 nm; and, these two sensitivities are uncorrelated across observers. Thus, the population behavior of $\Pi_1$ in the long-wavelength spectral region suggests that the primary and secondary modes are not due to one and the same pigment. Unfortunately, since the field sensitivity of $\Pi_1$ in the neighborhood of its secondary mode is at most a few percent of its peak sensitivity, and since the short-wavelength cones make negligible contribution to color matches to monochromatic light of wavelength $\lambda \geqq$ 540 nm or so (the trichromatic eye is dichromatic in this region of the spectrum), it is not possible to test directly the candidacy of $\Pi_1$ for a color matching fundamental in this crucial region of the spectrum. However, Pugh and Sigel (1978), in re-examining the fitting of linear combinations of color matching functions to the $\Pi$-mechanisms, have found some evidence that the spectral sensitivity of $\Pi_1$ in the region 500 nm $\leqq \lambda \leqq$ 540 nm is a cause of a material increase in the r.m.s. error of best fit. (Pugh and Sigel's fitting procedure minimized the sum of squared deviations of the logarithm of the linear combination of color matching functions from the log spectral sensitivity of the $\Pi$-mechanisms.) Nonetheless, the rejection of $\Pi_1$ as a fundamental cannot be considered categorical. In short, the spectral sensitivity alone of the $\Pi_1$ branch cannot strongly reject the hypothesis that it is determined by a single cone class.

*Field superposition experiments.* Fortunately, the hypothesis that the adaptation of the $\Pi_1$ branch is determined by a single cone class may be tested by means other than relating its spectrum to color matches, because of the laws that the color fundamentals must obey, viz. invariance and additivity. Now we know that to a reasonable approximation the $\Pi_1$ branch has a shape that is invariant across the spectrum (Stiles, 1939, Figs 37–39; Pugh, 1976, Fig. 4; see also Fig. 5 of this paper). Indeed, to say that the shape is invariant across the spectrum is simply to repeat that the branch obeys the Displacement Laws, which are necessary conditions for its having a unique spectral sensitivity function. One should not conclude, however, that the invariance of the shape of the $\Pi_1$ branch has been tested exhaustively. One may, however, begin an analysis of the $\Pi_1$ branch with the certain knowledge that it exhibits no clear-cut lack of invariance such as an inflection point on some fields and not on others. One can adopt the working assumption of an approximately invariant shape.

Given shape-invariance across the spectrum, the most powerfully diagnostic experiment to perform is a superposition experiment, for any mechanism whose adaptation state is determined by a single cone class must obey additivity. The assumption of field spectral invariance of the increment threshold curve, and the hypothesis of additivity, allow precise prediction of the effect on $\Pi_1$ of any mixture of any two monochromatic fields (Pugh, 1976). Figure 3A shows a test of field additivity of $\Pi_1$ for a mixture of 430 nm and

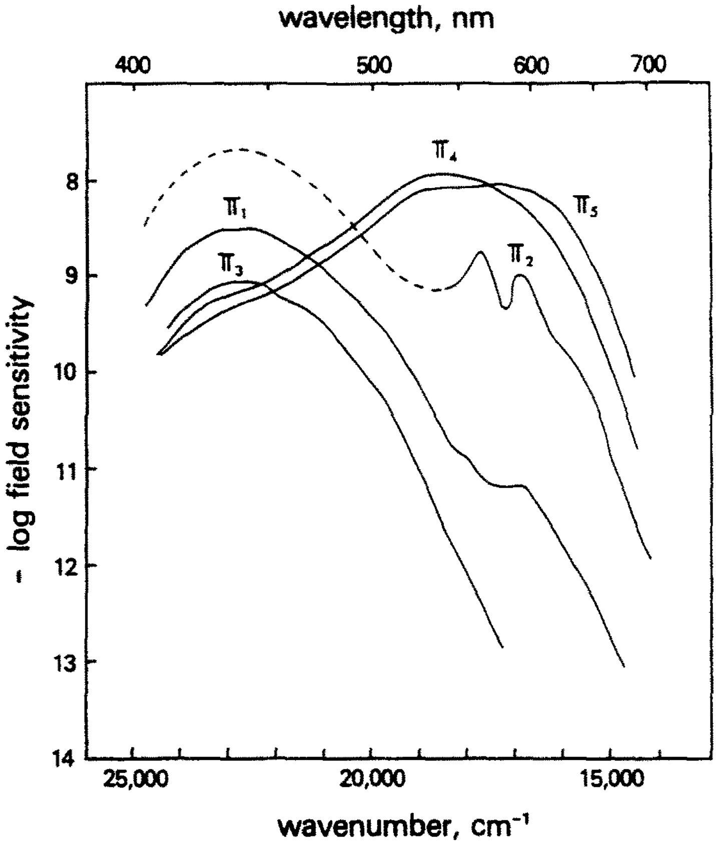


Fig. 2. The average field sensitivities of the four observers of Stiles (1953) for the five principal branches of the two-color threshold. Stiles (1978) gives details of the procedure used to obtain these curves.

500 nm fields, wavelengths which, though separated by 70 nm, are both in the neighborhood of the primary mode of the $\Pi_1$ field sensitivity (see Fig. 2). The results follow the prediction of additivity. Figure 3B shows a similar test performed with a mixture of 430 nm and 590 nm fields. Here, additivity fails systematically: because less of the mixture is required to produce a constant effect than the additivity hypothesis predicts, this failure is classified as a "super-additivity". Pugh (1976) showed that all mixtures of 430 nm fields with fields of 550 nm or greater result in similar super-additivity, if the field intensities $W_\mu$ are chosen so that $W_\mu \leqq 0.3/\Pi_{1\mu}$.

Another dramatic failure of additivity in $\Pi_1$ has recently been observed by Polden and Mollon (1979) in the parafovea. Figure 4 shows an example of their results. This "negative masking" or "cancellation of adaptation" is a sub-additivity, and at first sight incompatible with the super-additivity found by Pugh (1976). In fact, however, the intensity conditions for observing the effects are quite different. Super-additivity of bluish and yellowish fields is found when the fields are roughly $\Pi_1$-equated; sub-additivity or cancellation when the fields are roughly equated for color cancellation. The ratio of blue field/yellow field intensities is about 20 times greater for cancellation than for superadditivity.

The steady-state results just reviewed lead us to reject the hypothesis that the adaptation of the $\Pi_1$ branch is determined by a single cone class: for wavelengths $\mu \geqq 550$ nm cones other than the short-wavelength sensitive receptors must contribute signals that determine the state of $\Pi_1$ adaptation. If the concept of spectral invariance is extended into the dynamic domain, there is further evidence against the hypothesis that the $\Pi_1$ branch is determined by only one class of cones. Barring transient interference from other cone classes, any two lights *equated for their steady-state effect* on a single cone class should (1) "silently substitute" for one another and (2) result in identical time-courses of adaptation. Mollon and Polden (1975) demonstrated the failure of silent substitution for 520 nm and 580 nm $\Pi_1$-equated fields. Augenstein and Pugh (1977) showed that the time-courses of adaptation to and recovery from $\Pi_1$-equated fields are quite different in the short- ($\mu \leqq 500$ nm) and long-wavelength ($\mu \geqq 550$ nm) regions of the spectrum, (see Fig. 8). Thus, the evidence from dynamic experiments is consistent with that from the static experiments in showing that long-wavelength fields affect the state of adaptation of the $\Pi_1$ pathway through the mediation of signals from the middle- and/or long-wavelength sensitive cones.

## III. FORMAL THEORY OF $\Pi_1$ AND $\Pi_3$

*Static theory of $\Pi_1$*

Our initial premise is that an observer *operating on the $\Pi_1$ or $\Pi_3$ branch at threshold* detects a perturbation signal that is initiated exclusively by photons

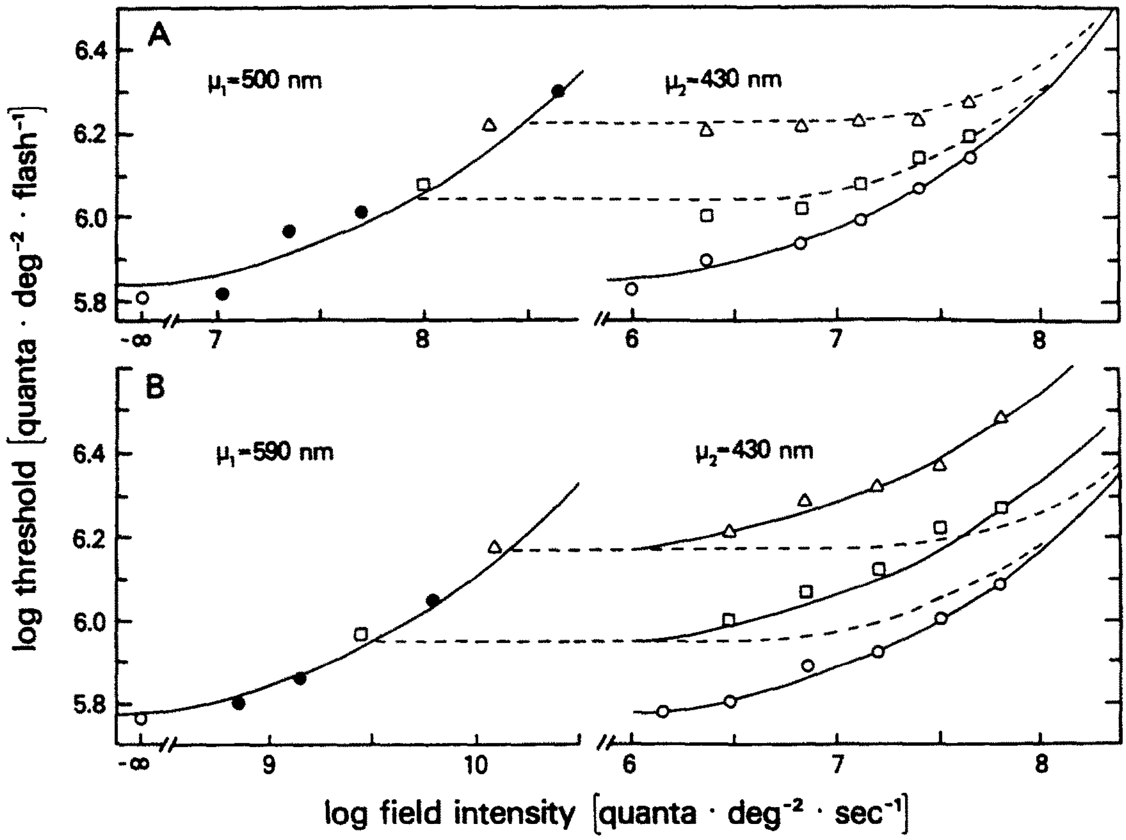


Fig. 3. A. A field-mixture experiment for the $\Pi_1$-branch, after Pugh (1976); one field component was $\mu_1 = 500$ nm; the other component was $\mu_2 = 430$ nm; the test was always a $\lambda = 435$ nm flash of 50 msec duration, presented foveally. The observer's threshold was allowed to reach steady-state on a 500 nm background of a given intensity, and then a series of increasingly intense 430 nm fields were admixed and the threshold was measured for each mixture. The solid line is the standard increment threshold function, which has been fitted by computer to the data obtained for each field component alone. The dotted lines are the predictions of the additivity hypothesis. Proof that the branch studied is the same as Stiles's $\Pi_1$ and details of the additivity prediction are given in Pugh (1976). The model of Eqns 1 and 2 predicts that additivity should obtain for this pair of fields. Observer, EP. B. A field mixture experiment for the $\Pi_1$ branch showing failure of additivity: $\mu_1 = 590$ nm; $\mu_2 = 430$ nm. The wavelength of the $\mu_1$-component was chosen because it is the region of the secondary mode (see Fig. 2) of the $\Pi_1$ field sensitivity curve. The standard increment threshold function (lower solid curves) was fit by computer to the $\mu_1$—or $\mu_2$—alone results. Again the dotted lines are the predictions of the additivity hypothesis. The model of Eqns 1 and 2 predicts that additivity should fail in the observed fashion. Indeed the solid curves are exactly those calculated with Eqns 1 and 2, providing we use the observer's field sensitivity to $\mu = 590$ nm on this day to estimate $[(K_2\beta_\mu)^n + (K_3\gamma_\mu)^n]^{1/n}$. The within-day estimate differs by about 0.25 log units from the mean across-day estimate. Observer SK.

absorbed in the short-wavelength sensitive or "blue" cones. For, under the conditions that isolate these two branches one observes a single test action spectrum or "tv$\lambda$" curve for $\lambda < 500$ nm or so. This spectrum has a pronounced peak at 430–440 nm, and falls over a log unit between 430 and 500 nm (Stiles, 1953, Fig. 14; Mollon and Polden, 1976, Fig. 3; Pugh, 1976, Fig. 2). The adaptive state of the long-wavelength sensitive mechanisms $\Pi_4$ and $\Pi_5$ can be varied enormously without affecting the relative test action spectrum in the short-wavelength spectral region. If it be allowed that these latter mechanisms ($\Pi_5$ and $\Pi_4$) represent detection of signals initiated predominately by the long- and middle-wavelength sensitive cones, and it also be allowed that the cones light adapt, then it follows that the middle- and/or long-wavelength cones make negligible contribution *to detection* under the conditions that isolate the $\Pi_1$ and $\Pi_3$ branches. Evidently, however, this "blue-cone" signal is attenuated by signals from the other cone classes when the threshold for the $\Pi_1$ branch is elevated by fields of wavelength $\mu \geq 550$ nm.

A mechanism by which the long- and middle-wavelength sensitive cones could effect $\Pi_1$ adaptation to long-wavelength fields was proposed by Pugh (1976) to account for the super-additivity of short- and long-wavelength fields: these cones send their signals to a "second site", effecting adaptation or signal attenuation at a locus in the pathway distinct from the "first site", whose state of adaptation is controlled exclusively by the activity of the "blue" cones themselves. Now, Polden and Mollon's (1979) cancellation results

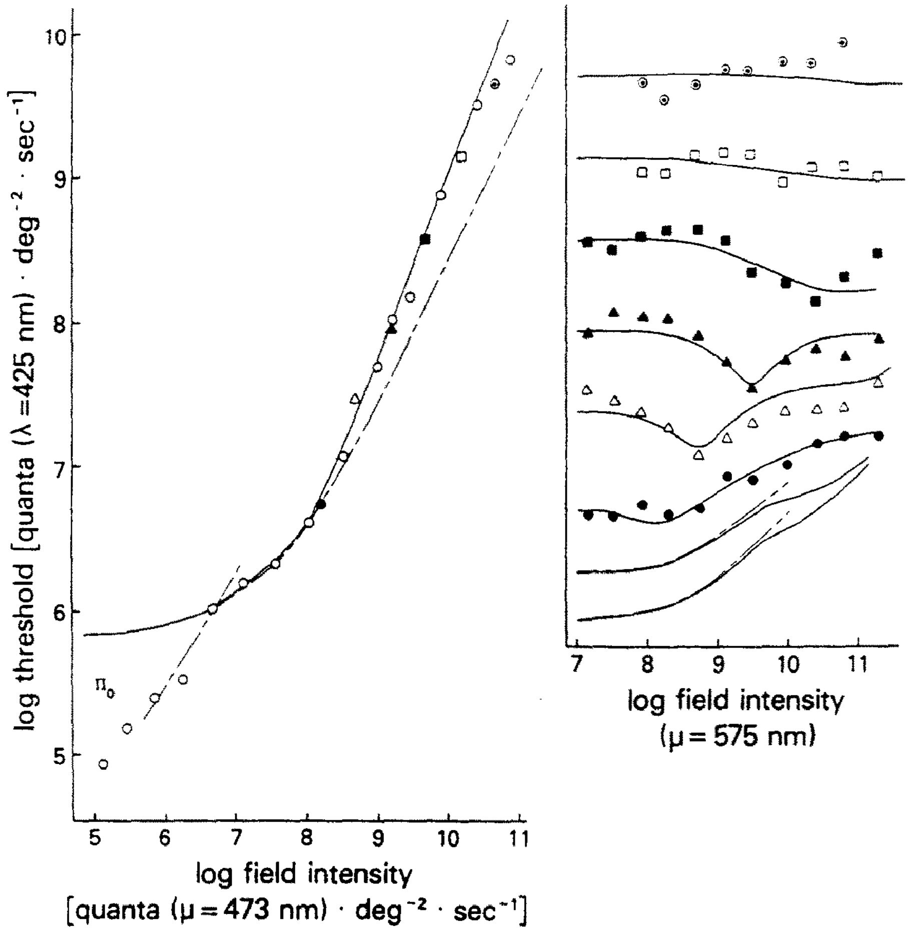


Fig. 4. A $\Pi_1$ field-mixture experiment in the parafovea, after Polden and Mollon (1979). The $\mu_1$-component of the mixture was a 473 nm field; at each of several fixed intensities of this field a series of 575 nm fields was admixed. Cancellation of $\Pi_1$ adaptation is observed for 473 nm fields of intensity greater than about $10^{8.8}$ quanta·deg$^{-2}$·sec$^{-1}$. The solid curves in both left-hand and right-hand panels were computed with the model of Eqns 1 and 2. The dashed curve in the left-hand panel (473 field alone) represents the Stiles template, fit to the lowest portion of the $\Pi_1$ branch. The dashed lines in the right-hand panel also represent the Stiles template, slid to agree with the theoretical curves (solid lines) generated by the model.

and those of Augenstein and Pugh (1977) require that the second site be chromatically opponent. Thus, we have the three essential elements of our static theory of $\Pi_1$: (1) a pathway for detection of short-wavelength flashes originating in the short-wavelength sensitive or "blue" cones; (2) a "first site" of adaptation controlled exclusively by the "blue" cones; (3) a "second site" of adaptation or attenuation controlled by the net steady-state signal of a "blue/yellow" chromatic input.

To put these three notions into a formal model we assume (a) that each site has a "gain" characteristic approximately like the standard $\zeta$-function of Stiles (Wyszecki and Stiles, 1967, p. 578), and (b) that the gain of the two-stage composite system is the product of the gain of its two components. Thus, letting $A$ represent an adapting field of arbitrary spectral composition, and $\alpha(A)$, $\beta(A)$, $\gamma(A)$ represent the quantum catch rates (see Appendix I) of the short-, middle- and long-wavelength sensitive cones, respectively, from $A$, then the composite gain is assumed to be given by

$$g(A) = \zeta_1[K_0\alpha(A)]\cdot\zeta_2\{|[K_1\alpha(A)]^n - [K_2\beta(A)]^n - [K_3\gamma(A)]^n|^{1/n}\} \quad (1)$$

where the relative threshold elevation of the pathway, and hence of the $\Pi_1$ branch, is given by

$$U_\lambda/(U_\lambda)_0 = 1/g(A). \quad (2)$$

To elucidate Eqn 1 for the purpose of the present discussion we may use the approximation $\zeta(x) \simeq 1/(1 + 9x)$. Then, combining Eqns 1 and 2 and taking logarithms, we obtain, as the threshold of the $\Pi_1$ branch in logarithmic units,

$$\log U_\lambda = \log(U_\lambda)_0 + \log[1 + 9K_0\alpha(A)] + \log\{1 + 9|[K_1\alpha(A)]^n - [K_2\beta(A)]^n - [K_3\gamma(A)]^n|^{1/n}\}. \quad (3)$$

Table 1. Parameters of best fit of two-site model, Eqns 1 and 2 to $\Pi_1$ field sensitivity curves[a]

| Observer | $-\log K_0$ | $-\log K_1^{b}$ | $-\log K_2$ | $-\log K_3$ | $n$ | RMS | RMS, $\mu \geq 526$ nm[c] |
|---|---|---|---|---|---|---|---|
| SK | 8.85 | 10.39 | 11.44 | 11.44 | 0.75 | 0.047 | 0.055 |
| EP | 8.80 | 10.95 | 12.04 | 11.67 | 0.71 | 0.028 | 0.041 |
| WS[d] | 8.84 | 10.66 | 11.66 | 11.58 | 0.82 | 0.041 | 0.040 |

[a] The cone action spectra used were $\Pi_3$, $\Pi_4$ and $\Pi_5$, normalized to unity. Pugh and Sigel (1978) show these to be excellent linear combinations of the small-field color-matching functions of Stiles and Burch (1958). Use of $\Pi_3$ and the Vos and Walraven (1971) R & G fundamentals instead yields equally good fits.

[b] The parameter $K_1$ is not well determined by increment threshold results below the $\Pi_3$ plateau. Rather than being treated as a free parameter, it was forced to satisfy $(K_1\alpha_{500})^n - (K_2\beta_{500})^n - (K_3\gamma_{500})^n = 0$, thus generating a "blue/yellow" cancellation equilibrium at $\mu = 500$ nm.

[c] The two-site model is most vulnerable in the spectral range 520–580 nm, where both sites are active below the $\Pi_3$ plateau. Thus, the RMS errors in this spectral region have been singled out.

[d] Stiles' average $\Pi_1$ observer (Wyszecki and Stiles, 1967). Data below 500 nm were obtained in the presence of an auxiliary field $\mu_\lambda = 555$ nm, $W_{\mu_\lambda} = 10^{9.3}$ quanta·deg$^{-2}$·sec$^{-1}$. (See Stiles, 1978, for details.)

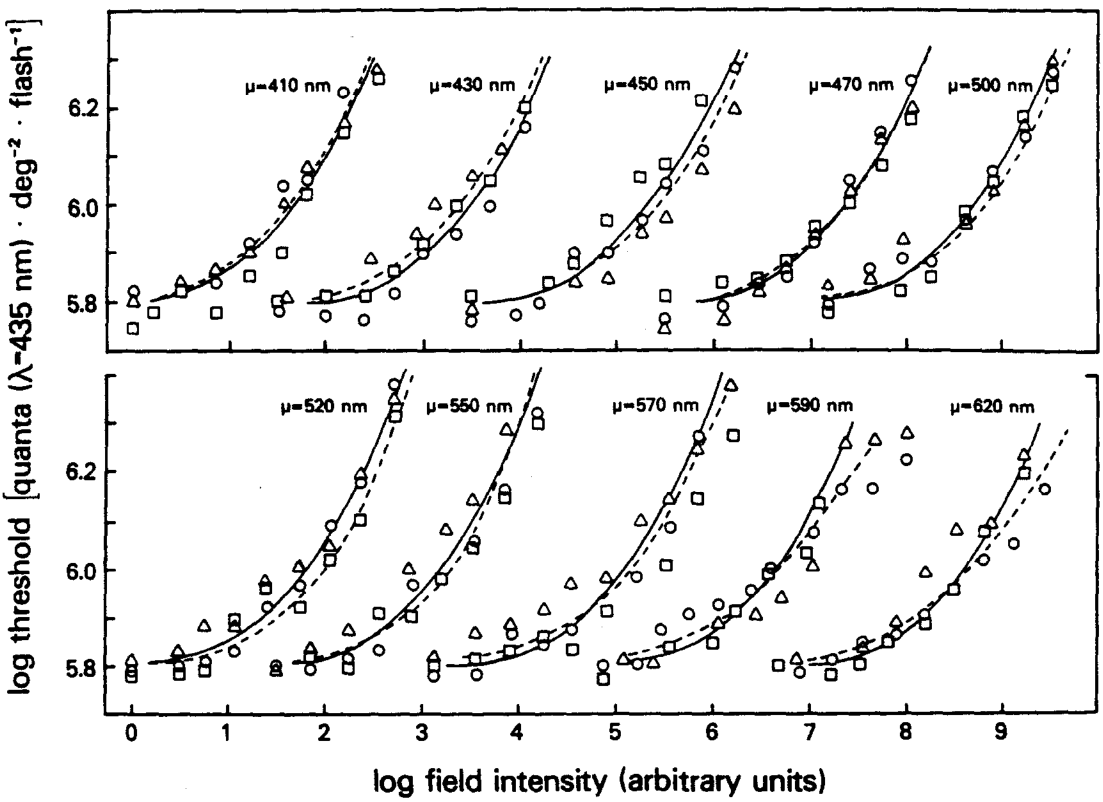


Fig. 5. Behavior of the theoretical model in the $\Pi_1$ range of the increment threshold (i.e. below the $\Pi_3$ plateau). The data are increment threshold curves obtained in 1974–1975 for observer EP. The test flash was a 1° foveal flash, $\lambda = 435$ nm, of 50 msec duration. A ~ 400 td 575 nm auxiliary field that did not elevate the $\Pi_1$ threshold was present continuously. In the investigation of Pugh (1976) each increment threshold curve was individually fit with the Stiles template and thus multiple estimates of the observer's field sensitivity were obtained for each field. The average of the individual estimates at each wavelength was taken as the observer's $\Pi_1$ field sensitivity. (See Pugh, 1976, Fig. 5.) The solid curve is the Stiles template, placed absolutely with respect to each set of data at approximately the position dictated by the observer's mean $\Pi_1$ field sensitivity. The dashed curves are the increment threshold curves generated by Eqns 1 and 2, instantiated with EP's parameter values (see Table 1); the dashed curves are exactly placed with respect to the solid curves, to which they were fit (Appendix II.)

The third term on the right-hand side in Eqn 3 embodies the color opponency of the second site. The power functions are used to give the expression some flexibility; however, there is no cogent reason other than simplicity to use this rather than any of a variety of other opponent functions as arguments for $\zeta_2$.

One *a priori* difficulty with Eqn 3 is that it predicts that increment threshold curves for the $\Pi_1$ branch will not obey absolutely the Field Displacement Law. To ascertain whether or not this prediction yields a material objection to the theory, and to instantiate the model for particular observers, we have done the following. We found, for the observers for whom a complete $\Pi_1$ field sensitivity is available, viz. Stiles's average observer (Wyszecki and Stiles, 1967, p. 579) and the two observers of Pugh (1976), the constants $K_0$, $K_1$, $K_2$, $K_3$ and $n$ that minimize approximately the squared error of the increment threshold curves computed with Eqn 1 (converted to an expression analogous to Eqn 3) from the "observed" $\Pi_1$ increment threshold curves for spectral fields. (See Appendix II for the rationale and further details of the fitting procedure.) Table 1 gives the values of the constants that yielded the best fits. Figure 5 shows some increment threshold curves for the $\Pi_1$ branch of observer EP obtained in 1974–1975 as part of the study of Pugh (1976). The solid curve is the Stiles template, approximately in the position of the observer's mean $\Pi_1$ field sensitivity. The dashed curves are the increment threshold curves calculated with the two-stage model of Eqn 1, parameters of Table 1, exactly placed with respect to the solid curve at each field wavelength.

These calculations demonstrate that a two-site model of $\Pi_1$ can generate reasonable increment threshold curves without seriously violating the Field Displacement Law, and do so in such a way as to recover to a reasonable degree the observer's $\Pi_1$ field action spectrum. The theoretical failure of the two-site model to generate shape-invariant increment threshold curves is not a material objection to it in fact: the discrepancies between the computed curves and Stiles's $\zeta$-function are within the error of measurements made so far. We may now turn our attention to the account that the model gives for the other steady-state phenomena associated with the $\Pi_1$ branch.

(a) *Super-additivity*. The two-stage model of $\Pi_1$ was developed (Pugh, 1976) precisely to account for the super-additivity of short- and long-wavelength fields that are roughly $\Pi_1$-equated. According to the model, long-wavelength fields ($\mu \geqq 570$ nm) of intensities $10^8$–$10^{10.5}$ quanta·deg$^{-2}$·sec$^{-1}$ (approx. 1000–20,000 td) affect only the second site; short-wavelength fields ($\mu \leqq 500$ nm) of $10^6$–$10^8$ quanta·deg$^{-2}$·sec$^{-1}$ significantly affect only the first site. Mixtures of such fields thus cause adaptation at both sites. For example, consider mixtures A of a $\mu_2 = 590$ nm field of intensity $W\mu_2 = 10^{10.1}$ quanta·deg$^{-2}$·sec$^{-1}$ with a series of $\mu_1 = 430$ nm fields of intensities $W\mu_1 = 10^6$–$W\mu_1 = 10^8$ quanta·deg$^{-2}$·sec$^{-1}$. This set of conditions is represented in the uppermost set of points (△) in the right-hand panel of Fig. 3B. For a given set of parameter values listed in Table 1 (e.g. SK's values, whose data are given in Fig. 3B) $K_0\alpha(A) = \Pi_{1430}W_{430}$; and $[K_1\alpha(A)]^n \ll [K_2\beta(A)]^n + [K_3\gamma(A)]^n$, so that $\{[K_2\beta(A)]^n + [K_3\gamma(A)]^n\}^{1/n} \simeq \Pi_{1590}W_{590}$, giving the observed result that mixing the 430 nm fields with the 590 nm field appears to slide the curve for 430 nm alone vertically rather than causing the points to follow the dashed line predicted by additivity.

(b) *Deviation of the increment threshold from the Weber line*. Another striking steady-state phenomenon associated with the "blue mechanism" is the deviation from the Weber line of the increment threshold on bright bluish backgrounds (Mollon and Polden, 1977a). The auxiliary field condition used by Mollon and Polden for studying this "saturation" effect is such that it isolates the $\Pi_3$ rather than the $\Pi_1$ branch; but we shall see below this is not a material difficulty, for we shall show how these two apparently distinct branches can be generated with the same two-site model. Indeed, even though the effect was first explicitly described by Mollon and Polden (1977a), deviation from the Weber line under conditions that isolate $\Pi_1$ ($\lambda = 420$ nm; $\mu = 435$ nm) can be found in published data of Stiles (1953, Fig. 11). Figure 4, left-hand panel, shows the effect. The uppermost points in the left-hand panel of Fig. 4 probably contains $\Pi_4$ intrusion (Polden and Mollon, 1979).

An upward deviation of the increment threshold from the Weber line is generated by the model when $[K_1\alpha(A)]^n$ is greater than $[K_2\beta(A)]^n + [K_3\gamma(A)]^n$, for under such conditions adaptation will increase concurrently at both sites as the intensity of $A$ increases. The solid line in the left-hand panel of Fig. 4, calculated with Eqns 1 and 2, shows how the model generates deviation from Weber-law behavior (see also Fig. 7, $\mu = 430$ nm). The model predicts a likely (though untested) difference between this "saturation" and the saturation of the rod increment threshold (Aguilar and Stiles, 1954), assuming the latter obeys "univariance" with respect to the action of the field, for the magnitude of the deviation from the Weber line generated by the model is not uniquely determined by the intensity of the main (bluish) field, but is rather dependent on the net signal at the second site, which is in turn dependent on the intensity of the continuously present "auxiliary field". This latter dependence has been demonstrated experimentally by Polden and Mollen (1979).

(c) *Cancellation of adaptation*. Cancellation of adaptation at the second site can be obtained in the model when the terms $[K_1\alpha(A)]^n$ and $[K_2\beta(A)]^n + [K_3\gamma(A)]^n$ are brought from a prior state of imbalance to a state of approximate equality. The effect can be seen most clearly if the adaptation state of the first site, determined by $K_0\alpha(A)$, is held approximately constant while the adaptation at the second site is cancelled: the experiment of Fig. 4 in which a series of increasingly intense long-wavelength fields is mixed with a relatively intense short-wavelength field achieves the goal. The solid curves in the right-hand panel of Fig. 4 have been calculated with Eqns 1 and 2. The parameter values used were $K_0 = 10^{-8.10}$, $K_1 = 10^{-9.92}$, $K_2 = K_3 = 10^{-10.5}$, $n = 0.82$. This value of log $K_0$ represents a sensitivity increase for the first site of about 0.7 log units over that in the fovea (see Table 1)—presumably due primarily to the lack of macular screening pigment in the parafovea. However, the value of $K_0/K_1$ is the same as that in Table 1 for the model fit to Stiles's average observer. The theoretical curves also require the use of a parameter $W_0$ that describes a "saturation" of the middle- and long-wavelength cones' steady-state signals to the second site, such us the saturation that would arise from the bleaching of the cone pigments. (See next section, and Appendix III). For the theoretical curves of Fig. 4 that describe PGP's parafoveal results, $W_0 = 10^{9.5}$ quanta·deg$^{-2}$·sec$^{-1}$, whereas $W_0 = 10^{10.5}$ seems to be the value required by the model to account for *foveal* results, including those from field-cancellation experiments like that reported in Fig. 4 (Pugh and Larimer, in preparation). It is also notable that the parameters $K_1$ and $K_2$ seem to be shifted about 1 log unit from the average values required for the fovea (Table 1).

*Extension of the theory to account for $\Pi_3$*

Up to this point we have avoided direct discussion of $\Pi_3$. We note immediately that the *relative* test spectral sensitivities of $\Pi_1$ and $\Pi_3$ appear identical in the spectral region $\lambda \leqq 500$ nm (Stiles, 1953, Fig. 14), the only region where one can be sure that the middle- and long-wavelength sensitive mechanisms are not contaminating the threshold measurements. And, indeed, the *relative* field sensitivities of $\Pi_1$ and $\Pi_3$ are identical in this same spectral region, and in perfect agreement with the test sensitivity. Since the spectral sensitivity in question declines over 1 log unit between 430 nm and 500 nm, one can be reasonably confident in asserting that the signals the observer detects at threshold on the $\Pi_1$ or $\Pi_3$ branch are initiated by photon absorptions in one and the same class of cones, the short-wavelength sensitive or "blue" cones.

The question to be asked then is this: what evidence is there that the $\Pi_3$ branch involves a visual pathway distinct from that responsible for $\Pi_1$? A glance at Figs 1 and 2 shows the evidence: (1) $\Pi_3$ has a distinct (absolute) threshold, about 0.5–0.7 log units above that of $\Pi_1$; (2) $\Pi_3$ has a field sensitivity distinct from that of $\Pi_1$, manifestly of different shape above 550 nm (though beginning to differ at about 500 nm), and differing in absolute magnitude (though not in relative sensitivity) by about $\frac{1}{2}$ log unit below 500 nm. Can the single-pathway, two-site model of Eqns 1 and 2 provide an account of these apparent distinctions?

Firstly, then, how might the apparent absolute threshold of the $\Pi_3$ branch, the plateau first known as the "limited conditioning effect" (Stiles, 1939), arise? Stiles's (1946a) study shows that the limited conditioning effect occurs at field intensities at which, as we now know from reflection densitometric studies, significant bleaching of chlorolabe and erythrolabe occurs. To be precise: in the nine observers (out of a sample of 20) that show the limited conditioning effect, the mean retinal illuminance of a 610 nm field that renders the observer $\frac{1}{2}$ log unit onto the plateau is 10,000 td, almost a 30% bleach. However, the value 10,000 td probably significantly underestimates the true population average, for insufficient field intensity is a likely reason that the effect is not found in the other 11 observers. (In 1939 Stiles observed the effect in his own eye, but in the 1946 study failed to do so for lack of adapting field intensity, (cf. Fig. 2 of Stiles, 1946a).

If the long-wavelength mode of the $\Pi_1$ field sensitivity is caused by *steady-state signals* to the "second site" from cones containing chlorolabe and erythrolabe, a plateau must occur, because these signals will have to approach an asymptote at, or somewhat before, the intensities at which significant bleaching occurs. To account formally for the effect of pigment bleaching we simply let the terms $\alpha(A)$, $\beta(A)$, $\gamma(A)$ in Eqn 1 represent *absorbed* quanta·deg$^{-2}$·sec$^{-1}$. Using Rushton's (1958) expression for the fraction pigment remaining at equilibrium, and letting $W_0$ represent the intensity in photons absorbed deg$^{-2}$·sec$^{-1}$ at which a 50% bleach occurs (and assuming equal $\lambda_{max}$ photosensitivities for the three cone pigments), one can readily show that as the intensity of an arbitrary light $A$ increases, $\gamma(A)$ and $\beta(A)$ and $\alpha(A)$ all approach the limit $W_0$, though at differential rates, depending on the spectrum of $A$ (Appendix III). Now, for a field of $\mu > 570$ nm or so $(K_1\alpha_\mu)^n \ll (K_2\beta_\mu)^n + (K_3\gamma_\mu)^n$, so that for long-wavelength lights the argument of $\zeta_2$ in Eqn 1 stabilizes at $(K_2^n + K_3^n)^{1/n} \cdot W_0$, and the model thus predicts an apparent "$\Pi_3$ plateau" with threshold

$$-\log \zeta[(K_2^n + K_3^n)^{1/n} \cdot W_0] \simeq \log[1 + 9(K_2^n + K_3^n)^{1/n} W_0]$$

—a threshold level that is not dependent on the wavelength of the field. The threshold of the pathway on such a long-wavelength field will rise above this plateau as the intensity becomes sufficient to cause adaptation at the first site. Figure 6 shows a series of increment threshold curves calculated with the two-stage model fitted to SK's $\Pi_1$ data (see Table 1), with allowance made for pigment bleaching. The half-bleaching constant used was $10^{10.50}$ quanta·deg$^{-2}$·sec$^{-1}$ (25,000 td at 555 nm). The solid curves in Fig. 6 are the standard Stiles increment threshold function slid for best agreement with the curves (symbols) computed with the model. In short, a quite reasonable modification of the model to include a saturation (such as would result from bleaching) of the middle- and long-wavelength cone input to the second site provides an explanation of the apparent $\Pi_3$ absolute threshold, i.e. of the "limited conditioning effect". (It should be noted that the signal saturation in discussion in this paragraph is a concept completely distinct from the deviation from the Weber line discussed previously. See Appendix III.)

Our second question concerns the apparent difference in *field sensitivity* between the $\Pi_1$ and $\Pi_3$ branches. The model accounts for the difference in $\Pi_1$ and $\Pi_3$ field sensitivities at long-wavelengths ($\mu \geqq 570$ nm or so; see Figs 1 and 6) as follows. For fields of such spectral composition the $\Pi_1$ branch represents changes in the adaptation state only at the second site; this is because below the $\Pi_3$ plateau such fields produce a negligible rate of photon absorptions in the "blue" cones [$K_0\alpha(A) < 0.001$]. At and immediately above the $\Pi_3$ plateau the signals from the long- and middle-wavelength cones remain constant at $\simeq (K_2^n + K_3^n)^{1/n} \cdot W_0$, while the term $(K_1\alpha)^n$ remains negligibly small: at these levels, then, changes in threshold depend exclusively on changes in the adaptation state

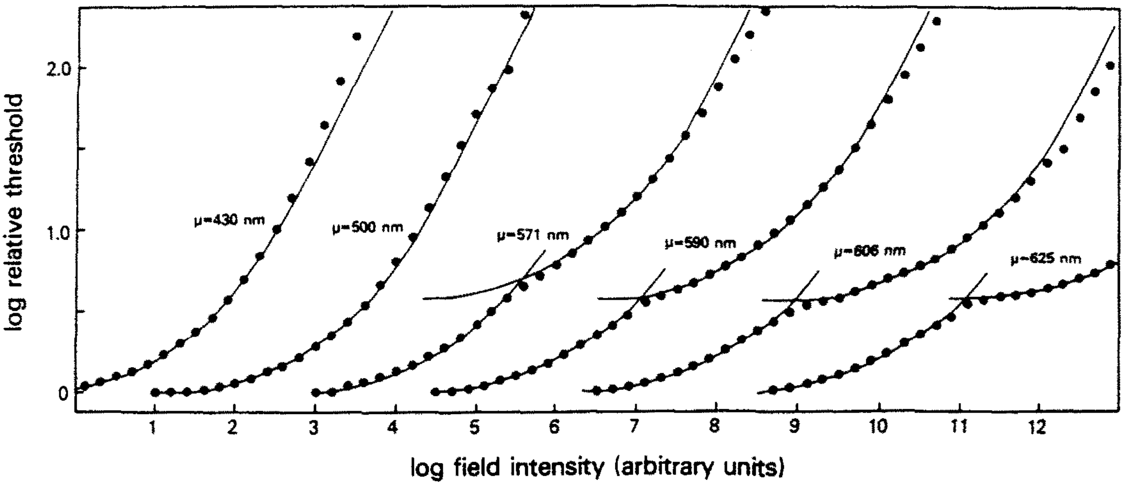


Fig. 6. A set of increment threshold curves (solid symbols) computed with the model of Eqns 1 and 2 with observer SK's parameter values (Table 1). The model has been extended into the $\Pi_3$ range by making allowance for pigment bleaching (see text). The smooth curves represent the standard Stiles template, slid for best fit to each "branch".

of the first site, controlled by $K_0\alpha(A)$. It is clear from Fig. 6 that in the long-wavelength region of the spectrum, where $(K_1\alpha)^n \ll (K_2\beta)^n + (K_3\gamma)^n$, the model generates differences in log field sensitivity between the apparent $\Pi_1$ and $\Pi_3$ branches that are of the correct order of magnitude. The model also yields a gradual diminution of this difference as one proceeds from longer to shorter wavelengths, with no $\Pi_3$ branch present at all at about 500 nm.

The remaining bit of evidence bearing on the possible difference in the pathway responsible for $\Pi_1$ and that responsible for $\Pi_3$ is the approximate $\frac{1}{4}$ log unit *absolute* difference in $\Pi_1$ and $\Pi_3$ field sensitivity below 500 nm (see Fig. 2; it should be noted that there is virtually no *relative* difference). To deal with this remaining issue, we must pay very close attention to the conditions used to isolate the $\Pi_1$ and $\Pi_3$ branches in the study from which the curves of Fig. 2 were derived (Stiles, 1953, 1978). Both experiments required the use of *auxiliary fields*. Our field mixture experiments (Pugh, 1976; Polden and Mollon, 1979) show that the role of these fields requires careful scrutiny. In order to isolate the $\Pi_1$ branch on short-wavelength fields, Stiles (1953, p.82) used a 555 nm auxiliary field of $10^{9.3}$ quanta$\cdot$deg$^{-2}\cdot$sec$^{-1}$; this field alone elevated the threshold of the branch about 0.2 log units (see Stiles, 1953, Fig. 10). To this auxiliary field were admixed the "main", short-wavelength fields, and the intensity of each main field required to obtain a criterion threshold elevation was estimated. As Stiles (1953, p. 82) points out, the auxiliary field procedure gives only the relative field sensitivity to the main field; in order to obtain the absolute field sensitivity one must make an assumption about how the effects of the auxiliary and main fields combine. It seems apparent that the assumption of field additivity was used in 1953 to obtain the absolute field sensitivities, an assumption now known to be invalid: 555 nm and short-wavelength ($\mu < 500$ nm) fields combine super-additively in their effects on the $\Pi_1$ branch (see Pugh, 1976, Fig. 10). It is not difficult to show that the assumption of additivity when "multiplicative" super-additivity obtains, leads under the stated auxiliary field condition to an over-estimation of the absolute field sensitivity below 500 nm of the $\Pi_1$ branch by about 0.25–0.35 log units. It is important to note that the relative field sensitivity is unaffected.

Once the effect of the auxiliary field used to isolate $\Pi_1$ is accounted for, there still remains unaccounted for a residual 0.15–0.25 log unit difference in the absolute field sensitivities of the $\Pi_1$ and $\Pi_3$ branches of Stiles's average observer to short-wavelength fields. Though this residual absolute difference is small, it cannot be discounted as random, owing to the close agreement of the relative $\Pi_1$ and $\Pi_3$ field sensitivities in the short wavelengths. To isolate $\Pi_3$ Stiles used a red auxiliary field that put the observer's threshold onto the plateau. In terms of the model, this would mean that the second site was adapted to the level determined by $(K_2^n + K_3^n)^{1/n}\cdot W_0$, the saturated signal from the middle- and long-wavelength cones, whereas the first site was unadapted. Any process that would lead to a loss of net signal at the second site during the course of the experiment would have led to a underestimation of the field sensitivity of the branch. One possibility is that the term $(K_1\alpha)^n$ becomes large enough to cancel part of the signal $(K_2^n + K_3^n)^{1/n}\cdot W_0$. This would require values of $-\log K_1$ about 1.0 log units lower than those in Table 1, if the model be taken strictly. Another possibility is that the extended exposure to an auxiliary field that maintains a >90% bleached state in the middle and long-wavelength cones actually leads to a loss of signal to the second site. Neither such effect would alter the relative field sensitivity estimated for the branch, but either would yield the result that the absolute field sensitivity of the first site was slightly under-estimated in the $\Pi_3$ isolation conditions.

In brief, then, we think that there is only a single absolute short-wavelength field sensitivity curve underlying the $\Pi_1$ and $\Pi_3$ results: this is the field sensitivity of the first site in the pathway, with absolute peak sensitivity of about $10^{-8.8}$ reciprocal quanta$\cdot$deg$^{-2}\cdot$sec$^{-1}$.

*Dynamic theory*

*Adaptation kinetics of the $\Pi_1$ pathway.* One of the most striking phenomena associated with adaptation under conditions that isolate the $\Pi_1$ and $\Pi_3$ branches of the two-color increment threshold is the large and relatively long-enduring transient rise in threshold at the extinction of long-wavelength ($\mu \geq 550$ nm) fields. This phenomenon, first reported by Stiles (1949a), has been studied in detail by Mollon and Polden (1976, 1977b) who have named it "transient tritanopia". The transient threshold elevation of the $\Pi_1/\Pi_3$ pathway can exceed the prior steady-state elevation by 1.5 log units; the time-constant of recovery of log threshold for the pathway can exceed 30 sec (Augenstein and Pugh, 1977). The offset transient threshold elevation is quite general: Mollon and Pollen (1977b) have demonstrated the effect (a) in the parafovea, (b) in the protanopic and deuteranopic eye, and (c) for modest field decrements, as well as at complete extinction of long-wavelength fields. Figure 7 shows the basic result, a comparison of the steady-state threshold of the $\Pi_1/\Pi_3$ branches on 580 nm fields with the thresholds obtained 400 msec after the extinction of the same fields. As was pointed out above and is shown in Fig. 8, short-wavelength ($\mu \leqq 500$ nm) fields, equated with long-wavelength fields for their steady-state effect on $\Pi_1$, do not cause "transient tritanopia" at their extinction: this failure of spectral invariance combines with the steady-state mixture experiments to reveal the composite nature of the cone signals controlling $\Pi_1$ adaptation.

Augenstein and Pugh (1977) obtained evidence that the peculiar dynamics associated with $\Pi_1$ adaptation result from events occurring at or proximal to the site in the pathway at which chromatic interaction occurs. They found, for example, that the time-course of $\Pi_1$ threshold recovery after a 5 min exposure to a 570 nm field of $10^{10.1}$ quanta$\cdot$deg$^{-2}\cdot$sec$^{-1}$ or to

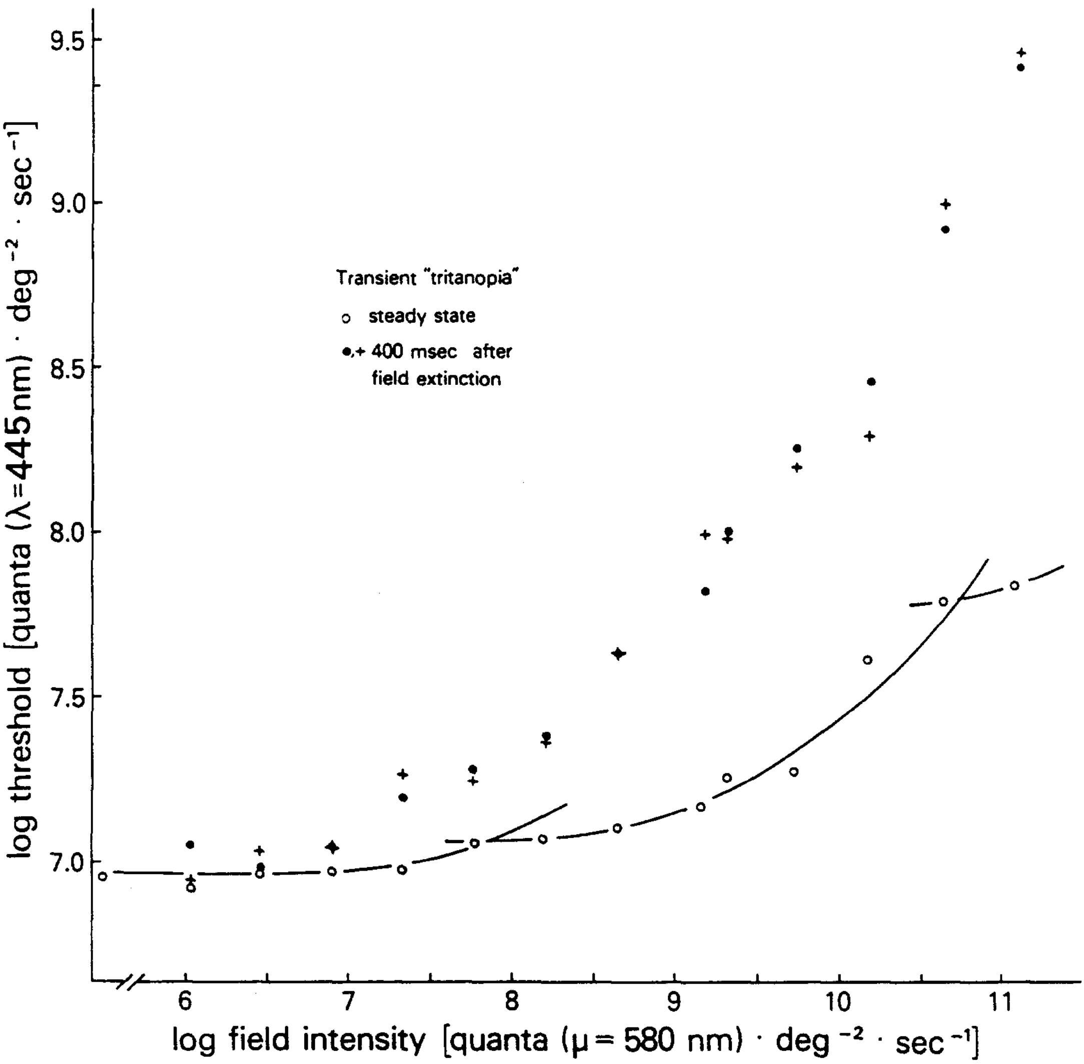


Fig. 7. O, steady-state increment thresholds for a foveally presented 1°, 200 msec, $\lambda = 445$ nm test flash in the presence of $\mu = 580$ nm fields of graded intensities; ●, + thresholds for the same flash 400 msec after extinction of the same fields. Observer, JM. From Mollon and Polden (1977b).

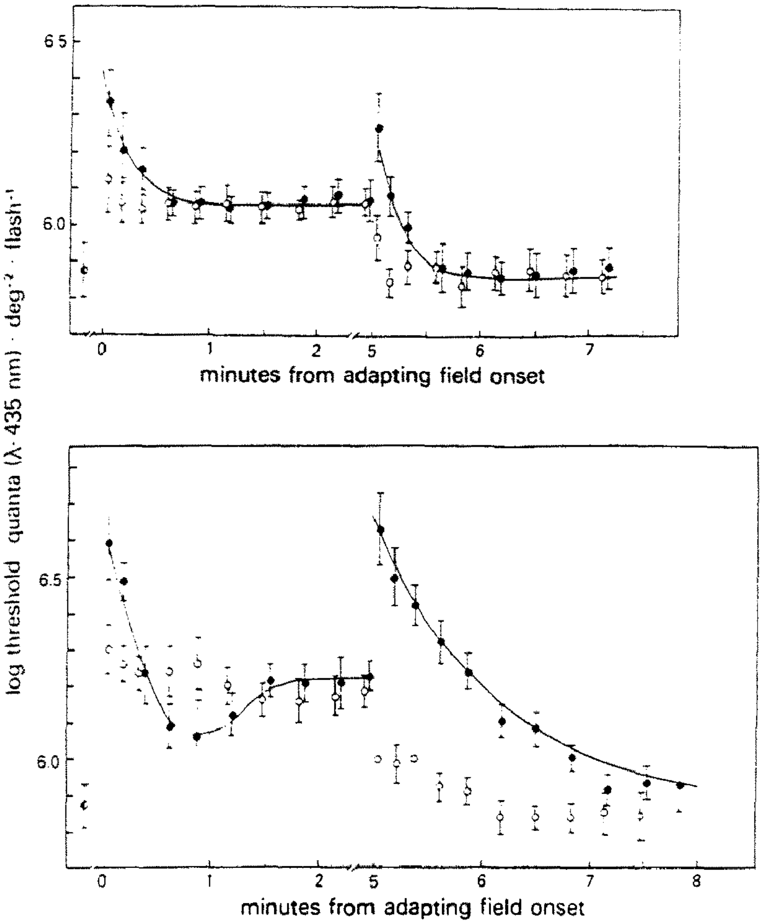


Fig. 8. Time-course of light adaptation to and recovery from 5 min exposures to $\Pi_1$-equivalent 430 nm and 590 nm fields, i.e. fields that effect the same elevation of the $\Pi_1$ threshold at steady state. Exposures and recoveries done in the continuous presence of 400 td yellow ($\simeq 570$ nm, $10^{8.7}$ quanta·deg$^{-2}$·sec$^{-1}$) "auxiliary" field which does not itself elevate the $\Pi_1$ threshold. Each set of points/bars is the mean $\pm$ 2 S.E.M. for five replications of the experiment; test flash 50 msec; $\lambda$ always 435 nm. The half-filled symbol represents the threshold during steady-state adaptation to the auxiliary field in the minute prior to the exposure of the main field. A. 430 nm field (open circles), flux density: $10^{7.43}$ quanta·deg$^{-2}$·sec$^{-1}$; 590 nm field (filled circles), $10^{9.44}$ quanta·deg$^{-2}$·sec$^{-1}$. Observer, SK. B. 430 nm field (open circles), flux density: $10^{7.74}$ quanta·deg$^{-2}$·sec$^{-1}$; 590 nm field (filled circles), $10^{10.22}$ quanta·deg$^{-2}$·sec$^{-1}$. Observer, EP. From Augenstein and Pugh (1977).

a 430 nm field of $10^{9.4}$ quanta·deg$^{-2}$·sec$^{-1}$ was slower than the recovery after a 5 min exposure to the approximately achromatic mixture of the same two fields. The 430 nm field must have cancelled part of the events causing transient tritanopia; but likewise, the 570 nm field must have cancelled some event(s) resulting in the sluggish recovery from the exposure to the 430 nm field alone. This mutual cancellation is the dynamic counterpart of the static cancellation effects (see Fig. 4) that have been investigated by Polden and Mollon (1979).

As may be seen in Fig. 8, the time-course of light-adaptation, as well as the course of "dark" adaptation under conditions that isolate the $\Pi_1$ branch, is quite different for short- and long-wavelength fields equated for their steady-state effect on the branch, a fact first pointed out by Stiles (1949b). Furthermore, as Fig. 8 shows, there is strong similarity between the time-courses for the light adaptation to and dark adaptation from long-wavelength fields. Because of this similarity, and because the very same long-wavelength fields that produce steady-state non-additivities when mixed with approximately $\Pi_1$-equated short-wavelength fields cause these "on" and "off" transients (Augenstein and Pugh, 1977), it seems reasonable to theorize that both "on" and "off" transients are produced by events occurring at one site, the same site posited to account for the steady-state super-additivity, cancellation and saturation effects. Our purpose here, then is to propose a simple formal model of this second or "opponent" site. In particular, it will be shown how the hypothesis of a slow restoring force, operative at the second site, can lead to a unified explanation of both on- and off-transients.

*Formalization.* Our first formalization is simply a way to represent the idea that the input to the second site is of a chromatically opponent nature. Let $V_\lambda(t)$, $V_\beta(t)$, $V_\gamma(t)$ represent time-dependent signals of the short-wavelength, middle-wavelength and long-wavelength sensitive cones, respectively, in response to an arbitrary light. Let $z = V_\alpha - V_\beta - V_\gamma$ and $f(z)$ be a monotone function satisfying:

$$f(z) < 0,\ z < 0$$
$$f(z) = 0,\ z = 0$$
$$f(z) > 0,\ z > 0.$$

Now let $V_2(t)$ represent the "polarization" of the second site, and $\sigma,\rho$ be positive constants satisfying $\sigma > \rho > 0$. We hypothesize, then, that $V_2$ satisfies the differential equation

$$\tau_2 \frac{dV_2}{dt} + V_2 = \sigma f(z) - \frac{\rho}{\tau_1}\int_0^t f[z(t')] \times \exp[-(t - t')/\tau_1]\,dt'. \quad (4)$$

In words, we hypothesize that the second site behaves as a low-pass ("$R$–$C$") filter that receives a feed-forward, subtractive signal convolved with a low-pass element of time constant $\tau_1$. This latter, feed-forward signal is our formal representation of the restoring

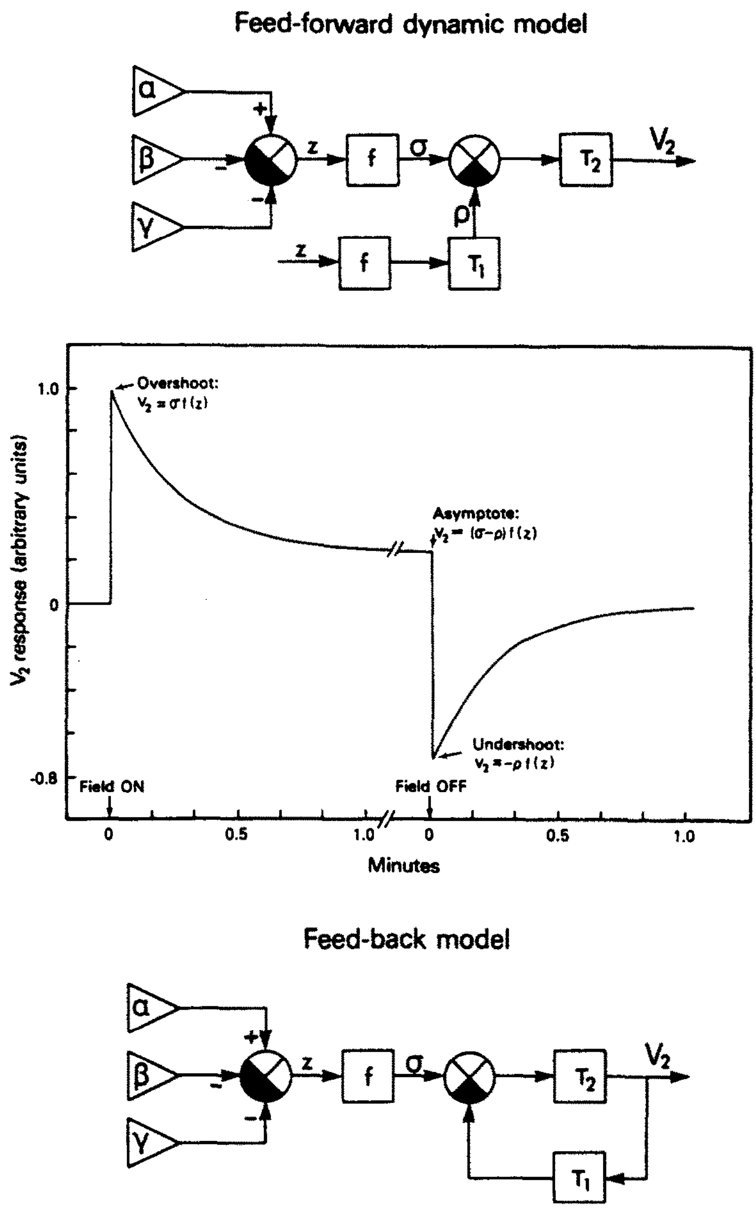


Fig. 9. A. Signal flow schematic of the feed-forward dynamic model (Eqn 4) of the $\Pi_1$ pathway. $z(t) = V_\alpha(t) - [V_\beta(t) + V_\gamma(t)]$ is the instantaneous difference between the short-wavelength and other cone signals. The square boxes represent "resistor–capacitor" type elements with the indicated time constants, $\tau_i$. $\sigma$ and $\rho$ are time-independent multipliers or gain factors that set the overall balance between the direct input $f[z(t)]$ to the "second site" and the delayed and inverted input. B. Step response of the model of 9A, assuming $f[z(t)] = \bar{f}$, a constant $t > 0$; "off-response" to the same step. C. Schematic for a feedback dynamic model of the $\Pi_1$ pathway that can behave like the feed-forward model if parameters are appropriately chosen.

force operating at the second site. The condition $\sigma > \rho$ simply constrains the asymptotic response to a positive input $z$ to be positive. Figure 9A shows a schematic diagram of the model.

*Step response to a long-wavelength field, $\mu \geqq 570$ nm.* According to our hypothesis, $\Pi_1$ adaptation to a step of light of intensity $W_\mu$ and wavelength $\mu \geqq 570$ nm is determined solely by signals from the middle- and/or long-wavelength cones to the second site; for such a step then, $z = V_\beta + V_\gamma$. Because the spectral sensitivity of $V_2$ in response to such a field $W_\mu$ must be that of $\Pi_1$, $z = V_\beta + V_\gamma = K(\Pi_{1\mu} W_\mu)^n$, where $K$ is a positive constant. Furthermore, since the time-scale of the results to be described (Fig. 8) is tens of seconds, and the response of the cones to steps of light may be expected to stabilize in fractions of seconds to seconds, we may treat $V_\beta$ and $V_\gamma$ as instantly determined by $W_\mu$. Let $f(V_\beta + V_\gamma) = f(z)$ be the input to $V_2$ for such a step $W_\mu$. The solution to Eqn 4, subject to $V_2(0) = 0$, is

$$V_2(t) = (\sigma - \rho)f(z)[1 - \exp(-t/\tau_2)] + \frac{\rho\tau_1}{(\tau_1 - \tau_2)} f(z)[\exp(-t/\tau_1) - \exp(-t/\tau_2)]. \quad (5)$$

If the restoring force builds up relatively slowly, i.e. if $\tau_1 > \tau_2$, there will be a transient overshoot of $V_2$ beyond the asymptote $(\sigma - \rho)f(z)$. For a given $\sigma,\rho$ a maximum overshoot is attained if $\tau_1 \gg \tau_2$, and this maximum approaches $\sigma f(z)$.

The relaxation of $V_2$ from a steady-state polarization $V_2(\infty) = (\sigma - \rho)f(z)$ to $V_2 = 0$ may also be readily obtained:

$$V_2(t) = (\sigma - \rho)f(z)\exp(-t/\tau_2) - \frac{\rho\tau_1}{(\tau_1 - \tau_2)} f(z)[\exp(-t/\tau_1) - \exp(-t/\tau_2)]. \quad (6)$$

As in the case of the on-response to a step, the off-response will have a transient, an undershoot below zero. For a given $\rho,\sigma$ this undershoot approaches $-\rho f(z)$ for $\tau_1 \gg \tau_2$. Figure 9B shows a plot of Eqns 5 and 6 for $\tau_1 = 15$ sec, $\tau_2 = 0.1$ sec and $\rho = 0.75\ \sigma$.

*Time-course of threshold recovery from a long-wavelength field.* Our two-site theory specifies that under the conditions that isolate the $\Pi_1$ pathway the signals detected at threshold originate in the short-wavelength sensitive cones' perturbation responses to the brief, short-wavelength test flashes. To complete our model and link Eqns 5 and 6 to the time-course of the threshold, we assume that threshold is attained when an impulse of intensity $\bar{U}_\lambda$ delivered at time $t$ causes a perturbation $\Delta V_2(t)$, the maximum of whose absolute value exceeds $|V_2(t)|$ by a constant ratio[4]:

$$\text{const} = \frac{|\Delta V_{2,\max}|}{|V_2(t)| + \theta}. \quad (7)$$

(The parameter $\theta$ is a constant that specifies the "noise" above which the perturbation must be detected when $V_2 = 0$.) Let $\alpha_\lambda$ be the short-wavelength cones' spectral sensitivity and assume that the cone perturbation response is brief relative to the time scale of events of interest (tens of seconds). The maximum absolute perturbation of $V_2$ caused by the test flash can be shown to be proportional to $f'(z)K_1\alpha_\lambda\bar{U}_\lambda$. Thus, Eqn 7 becomes

$$\text{const} = \frac{f'(z)K_1\alpha_\lambda\bar{U}_\lambda}{f(z)|\bar{V}_2(t)| + \theta} \quad (7')$$

where $\bar{V}_2(t)$ is here the response of $V_2$ to the unit step, $f(z) \equiv 1$, $z > 0$. Equation 7′ may be used to determine the function $f$, because the steady-state threshold must recover the shape of the increment threshold curve. Equation 7′ specifies the time-course of $\Pi_1$ adaptation to and recovery from long-wavelength fields of arbitrary intensity within the $\Pi_1$ range. In particular, it can be shown that the time-course of threshold at extinction of a steady-state long-wavelength field $W_\mu$ is given by

$$\pi_{1\lambda}U_\lambda = 1 + C(x)\left|(\sigma - \rho)\exp(-t/\tau_2) - \frac{\rho\tau_1}{(\tau_1 - \tau_2)}[\exp(-t/\tau_1) - \exp(-t/\tau_2)]\right| \quad (8)$$

where $C(x)$ is a function of $x = \Pi_{1\mu}W_\mu$ specified up to a multiplicative constant. Figure 10 shows a series of threshold recovery curves calculated with Eqn 8: the heavy solid curve is the standard increment threshold function $-\log \zeta(x)$, and gives the steady-state threshold prior to field extinction. The solid theoretical recovery has a restoring force time-constant $\tau_1 = 35$ sec; for the dashed curve $\tau_1 = 10$ sec. In all cases $\tau_2 = 0.08$ sec and $\rho = 0.5\ \sigma$. The data plotted are taken from Fig. 8. They were selected because they represent well *the extremes* of recovery times observed by Augenstein and Pugh (1977).

The value of $\tau_1$, the restoring force time-constant, may be estimated without computing the entire recovery time course, providing $\tau_2 < 0.10$ sec or so. For then Eqn 8 becomes

$$\log \pi_{1\lambda}U_\lambda = \log[1 + C\exp(-t/\tau_1)]. \quad (9)$$

Augenstein and Pugh (1977) used Eqn 9 to estimate $\tau_1$ from 18 (mean) recovery curves (about 90 experiments). For one observer the estimate $\pm$ 2 S.E.M. was $15.5 \pm 2.1$ sec, significantly smaller than that of the other observer, $\tau_1 = 27.6 \pm 6.6$ sec. One interesting qualitative prediction of the theory is that observers with stronger restoring forces should have lower long-wavelength $\Pi_1$ field-sensitivity, and greater magnitude off-transients. This prediction appears to be born out in the limited results now available (Augenstein and Pugh, 1977).

*Absence of the offset transient*

Another interesting phenomenon associated with

[4] The criterion of exceeding $|V_2(t)|$ may be generalized to exceeding any linear transformation (such as averaging) of $|V_2(t)|$ over some relatively brief time-interval (e.g. 1 sec), without change of results. The time-interval must be brief relative to the time-scale of the recovery measured, and of the same order as the perturbation response to the test flash.

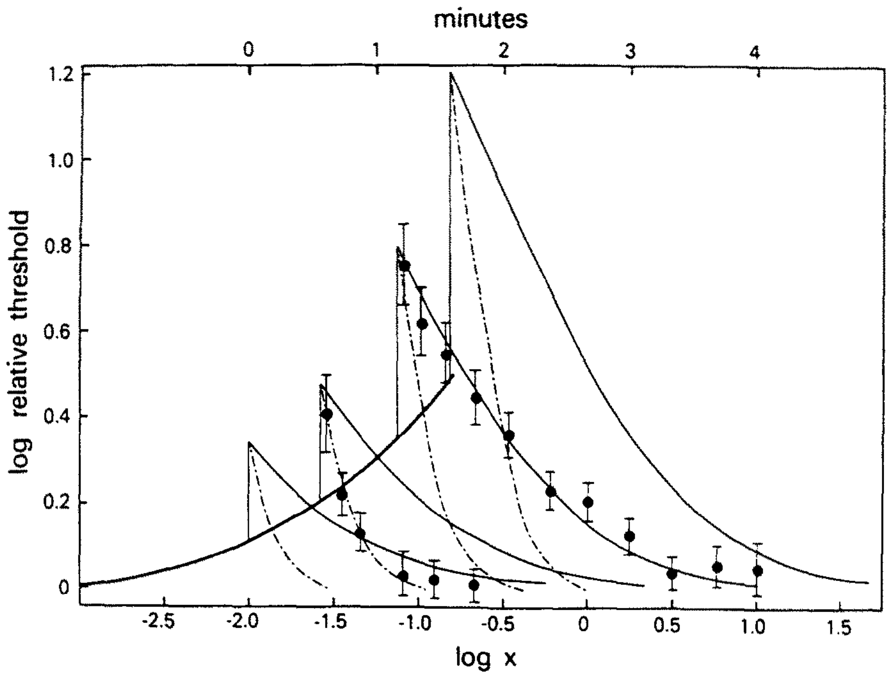


Fig. 10. Use of the feed-forward model to describe the recovery of threshold following the extinction of a field of a wavelength and intensity such that it adapts the $\Pi_1$ pathway only at the second site. The heavy solid curve is the standard increment threshold curve, $-\log \zeta(x)$, plotted against the lower abscissa, $\log x = \log \Pi_\mu W_\mu$. The solid and dashed lines represent the time-course of threshold recovery after steady-state adaptation to fields of various strengths, the value of the steady-state threshold before field extinction being indicated by the point on the standard curve from which the recovery curve departs. The upper abscissa gives the time scale. The solid theoretical recovery curves are calculated from Eqn 8, with a restoring force time-constant $\tau_1 = 35$ sec: other parameters are $\tau_2 = 0.08$ sec, $\rho/(\sigma - \rho) = 1.0$. The function $C(x)$ has been estimated by requiring agreement of Eqn 8 with the threshold level prior to the extinction of the field. The dashed recovery curves have restoring force time constant $\tau_1 = 10$ sec and all other parameters identical to those that generate the solid lines. The data plotted are taken from the results in Fig. 8, and represent well the range of recovery times observed. SK's recoveries are systematically faster than those of EP at all intensities of the pre-adapting field (see text).

the offset transient is its precipitous decline with intense adapting fields (Mollon and Polden, 1976, 1977b). For example, in the study of Mollon and Polden (1977b), for four normal observers the peak magnitude threshold elevation for a 445 nm test flashed 400 msec after the extinction of a 575 nm field occurs at approximately $10^5$ td; for a field of $10^{5.5}$td, the offset transient appears absent. An attractive explanation in terms of the general two-site model is this: above $10^5$ td the long- and middle-wavelength cone signals to the second site remain constant (see p. 308); the short-wavelength cone signal to the second site will continue to grow in magnitude, however, with increasing field intensity, gradually cancelling the constant signal from the other cones. This explanation is inconsistent with the quantitative version of the model we have presented, i.e. with Eqns 1–4 and the parameter values of Table 1: the values of $-\log K_1$ in the table are too large. For example, the model using SK's parameter values predicts that the peak magnitude of the offset transient should occur at $10^{5.7}$ td at 571 nm, and that complete abolition of the effect should not occur until $10^{7.7}$ td. It should be noted, however, that the prediction of the intensity at which the peak effect occurs depends critically on the spectral sensitivity of the short-wavelength cones at long-wavelengths.

Another possible explanation of the abolition of the offset transient is a gradual loss of the "restoring force" under lengthy exposure to fields that bleach most of the long- and middle-wavelength cone pigments. Determination of the action spectrum of the abolition effect in the region, say, 550 nm to 590 nm, should readily decide between a second-site cancellation due to a short-wavelength cone signal and a deteoriation of the restoring force: the former effect should have a steep decline in spectral sensitivity in the stated region; the latter effect, almost none.

## IV. DISCUSSION

### *The nature of the theory*

A solid body of experimental results contradicts the hypothesis that the light adaptation process or processes revealed by the $\Pi_1$ branch of the two-color threshold is controlled by a single class of cones. We have proposed here an alternative hypothesis, one that is capable at present of giving a unified account of a number of disparate results. The essence of the theory is the notion that thresholds of the $\Pi_1$ and

$\Pi_3$ branches manifest two types of "adaptation" events occurring in a single visual pathway, a pathway originating in the short-wavelength cones: the events of the second type occur proximal to the convergence of some antagonistic signals from the middle- and long-wavelength cones into this pathway. We emphasize that the notions "cone", "pathway", "site", etc. used to state the general theory are meant to have precise functional significance (Appendix IB).

We have instantiated our general notions in Eqns 1 and 2 (static theory) and in Eqns 4 and 7 (dynamic theory). While representing the theory in these specific equations we do not intend them to stand as uniquely correct representations of the general notions. Rather our purpose is to demonstrate that one relatively simple mathematical representation of our hypothesis can exhibit the variety of requisite qualitative behavior (superadditivity, cancellation, "saturation", the $\Pi_3$ plateau, etc.), and indeed produce first-order quantitative agreement with experimental results. We have no doubt, however, that the present representation involves some oversimplifications and only weakly supported assumptions. For example, use of the power-law expressions in Eqn 1 is tantamount to assuming that $V_\alpha(t)$, $V_\beta(t)$, $V_\gamma(t)$—the relevant measures of all three cones' responses to a light $A$—converge asymptotically in time to $V_\alpha(\infty) = \alpha(A)^n$, etc. where the power $n$ is the same. This power law expression was introduced primarily to allow some additional leeway in fitting the $\Pi_1$ increment threshold curves (Appendix II): it leads to computational simplicity, but is not strongly determined by the results.

Similarly, and to an even greater degree, the dynamic model (Eqns 4 and 7) involves several arbitrary choices and simplifications. It is possible, for example, to represent the essential notion of the restoring force in terms of a "feedback" model: a schematic of such a model is given in Fig. 9C. This feedback model can exhibit the same variety of behavior as (and indeed, a somewhat wider variety than) the feed-forward model (Fig. 9A). For $\tau_1 \gg \tau_2$ and appropriate choice of the remaining parameters, the models are indistinguishable. On the other hand, the feedback version gives a restoring-force time-constant (time-constant of recovery from the on- and off-transients) of $\tau_1/(1 - \rho)$, and so predicts that observers with the greater magnitude transients should have faster recovery rates—other things being equal. Neither the feed-forward nor feed-back models provides an excellent account of the on-transients, in part because of the assumption that the cone responses to the adapting fields stabilize instantly.

*Opponent colors theory*

Opponent colors theory (Hurvich and Jameson, 1957) postulates the existence of a visual pathway that codes the mutually exclusive sensations of blueness and yellowness. In particular, to account for the quantitative results of the blue/yellow cancellation experiments (Jameson and Hurvich, 1955) and for the invariance and approximate closure laws obeyed by blue/yellow equilibrium lights (Krantz, 1975b; Larimer, Krantz and Cicerone, 1975), opponent colors theory postulates a site in the visual system that receives signals of one sign from the short-wavelength sensitive cones and signals of opposite sign from the other cone classes, signals that can thus cancel one another's effects on the opponent pathway. One is tempted to identify the hypothetical site of cancellation involved in our experiments and that of opponent colors theory. Indeed, Mollon and Krauskopf (1973) put forth the hypothesis that many of the anomalous properties attributed to the "blue cones" might well be due to an anatomy in which these cones could send their signals only along "chromatic channels". Transient tritanopia, for example, involves the detection of a bluish test flash in the presence of a strong bluish after-image. Identification of the two hypothetical sites can be made only by quantitative comparisons (within individual observers) of the two types of experiments involved, and we feel that considerable circumspection is still in order.

*Relationship to other threshold studies*

The key empirical results—beyond the discovery of the $\Pi_1$ and $\Pi_3$ branches and measurement of their test and field sensitivities by Stiles—upon which our theory rests have been obtained in field-mixture experiments (Pugh, 1976; Augenstein and Pugh, 1977; Polden and Mollon, 1979). In this type of experiment the test is chosen so that at threshold the observer is operating on a single specified branch or mechanism, and the effects of various mixtures of two adapting fields on the one branch are measured. Prior to the studies just cited there was no field mixture experiment on the $\Pi_1$ and/or $\Pi_3$ branches in the literature, though Brindley (1970, p. 257) had suggested the potential value of such an experiment.

However, a number of test-mixture experiments involving the $\Pi_1$ and $\Pi_3$ mechanisms have preceded our work (Boynton, Ikeda and Stiles, 1964; Stiles, 1967; Ikeda, Uetsuki and Stiles, 1970; Krauskopf, 1974). In contrast with the field-mixture experiment, in a typical test-mixture experiment two or more mechanisms often are simultaneously active in determining threshold. The results of Boynton *et al.* (1964), Stiles (1967) and Krauskopf (1974) support the conclusion that signals from the short-wavelength and other cone classes are not probabilistically independent in their contributions to the detection of some composite flashes, and indeed can interact inhibitorily to a degree. The work of Guth (1965, 1967), though not explicitly dealing with specific $\Pi$-mechanisms, clearly shows cancellative interactions between pairs of short- and middle- or long-wavelength test flashes at threshold. These test-mixture experiments, like the opponent hue cancellation experiments that preceded them (Jameson and Hurvich, 1955) are precursors of our work in that they argue for the existence of a site or sites to which oppositely signed signals from the short-wavelength and other cone classes converge. However, there is no reason at present other than parsimony to assume that the site of such test interactions is the same as that we hypothesize to account for our field cancellation results. Our theory would need a fair amount of elaboration to predict the conditions under which test-flash interactions would occur.

*Anatomical sites*

The proper understanding of the notions "first site" and "second site" of adaptation proposed here lies in the interpretation of the mathematical formulations, and not in speculation about the retina. Nonetheless, some interesting possibilities bear brief mention.

It is now well-established that vertebrate cone photoreceptors light adapt, changing their speed and sensitivity (Baylor and Hodgkin, 1974; Norman and Werblin, 1973; Hood and Hock, 1975). Assuming the cross-sectional absorption area of a short-wavelength sensitive human foveal cone to be about $10^{-5}$ $deg^2$ of visual angle, the transmissivity of optical media at 430 nm to be 10%, and the *in situ* photopigment optical density to be 0.5, then a 430 nm field of $10^{8.8}$ quanta$\cdot$deg$^{-2}\cdot$sec$^{-1}$ should result in about 430 quanta absorbed/cone$\cdot$sec. Thus a field that, in the theory, gives rise to a 1 log unit threshold elevation at the first site is well within the range of intensities known to cause significant change in the time-scale and sensitivity of the turtle cones studied by Baylor and Hodgkin (1974). Given the 50–70 times greater volume of turtle cones, it seems reasonable to believe that fields which we have hypothesized to effect adaptation at the "first site" in the $\Pi_1$ pathway have indeed caused adaptation in the receptors themselves.

Electrophysiological research has demonstrated opponent recoding of receptor signals at the ganglion cell level of the primate retina (de Monasterio, Gouras and Tolhurst, 1975a, b) and thus one might be led to speculate that the cancellation phenomena occur at that retinal layer. However, recently Valeton and van Norren (1979) have demonstrated that "transient tritanopia" can be observed in the *b*-wave of the primate electroretinogram, under exactly the same stimulus conditions that give rise to the psychophysically measured phenomenon in man. When coupled with the hypothesis that the *b*-wave is generated in the inner nuclear layer, Valeton and van Norren's results argue that the chromatic interactions described in our theory occur either there or in the outer plexiform. Even if a literal physiological interpretation of the two-site theory be attempted, however, there is no reason to suppose that "adaptation" at the second site results from the same physiological process as that which causes "adaptation" at the first site.

*Conclusion*

Perhaps the greatest value in stating a theory such as that proposed here is the guide it affords to future experimentation. We think that any quantitative successes of the theory at present should be given less emphasis than qualitative predictions, because we have no knowledge of what alternative hypotheses could yield equally good fits.

Some quantitative aspects of the theory nonetheless deserve attention. The photon flux levels at which "first-site" and "second-site" adaptation events occur can be precisely determined; knowledge of these levels should play an important role in further development of color theory.

*Acknowledgements*—This work was supported in part by NSF grant BNS 76-14192 to E. Pugh and by MRC grant G976/139/N to J. Mollon. We thank M. Alpern, T. Adelson, J. Krauskopf, J. Larimer, D. MacLeod and B. Wandell for helpful comments on previous drafts. We thank J. Nachmias and NSF grant 75-27855 AO for computer time.

## REFERENCES

Aguilar M. and Stiles W. S. (1954) Saturation of the rod mechanism of the retina at high levels of stimulation. *Optica Acta* **1**, 59–65.

Alpern M. and Pugh E. N. Jr (1974) The density and photosensitivity of human rhodopsin in the living retina. *J. Physiol.* **237**, 341–370.

Alpern M., Rushton W. A. H. and Torii S. (1970) Signals from cones. *J. Physiol.* **207**, 463–475.

Augenstein E. J. and Pugh E. N. Jr (1977) The dynamics of the $\Pi_1$ colour mechanism: further evidence for two sites of adaptation. *J. Physiol.* **272**, 247–281.

Baylor D. A. and Hodgkin A. L. (1974) Changes in time scale and sensitivity in turtle photoreceptors. *J. Physiol.* **242**, 729–758.

Baylor D. A., Hodgkin A. L. and Lamb T. D. (1974) Reconstruction of the electrical responses of turtle cones to flashes and steps of light. *J. Physiol.* **242**, 759–791.

Bowmaker J. K., Dartnall H. J. A., Lythgoe J. N. and Mollon J. D. (1978) The visual pigments of rods and cones in the Rhesus monkey, *Macaca mulatta*. *J. Physiol.* **274**, 329–348.

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. (1957) Two theorems in color vision. *Q. J. exp. Psychol.* **9**, 101–112.

Brindley G. (1970) *Physiology of the Retina and Visual Pathway*, 2nd edn. Edward Arnold, London.

Guth S. L. (1965) Luminance addition: general considerations and some results at foveal threshold. *J. opt. Soc. Am.* **55**, 718–722.

Guth S. L. (1967) Nonadditivity and inhibition among chromatic luminances at threshold. *Vision Res.* **7**, 319–328.

Hood D. S. and Hock P. A. (1975) Light adaptation of the receptors: increment threshold functions for the frog's rods and cones. *Vision Res.* **15**, 545–554.

Hurvich L. M. and Jameson D. (1957) An opponent-process theory of color vision. *Psychol. Rev.* **64**, 384–404.

Ikeda M., Uetsuki T. and Stiles W. S. (1970) Interrelations among Stiles $\Pi$ mechanisms. *J. opt. Soc. Am.* **60**, 406–415.

Jameson D. and Hurvich L. M. (1955) Some quantitative aspects of opponent-colors theory. I. Chromatic responses and spectral saturation. *J. opt. Soc. Am.* **45**, 546–552.

King-Smith P. E. (1973a) The optical density of erythrolabe determined by retinal densitometry using the self-screening method. *J. Physiol.* **230**, 535–549.

King-Smith P. E. (1973b) The optical density of erythrolabe determined by a new method. *J. Physiol.* **230**, 551–560.

Krantz D. H. (1975a) Color measurement and color theory: I. Representation theorem for Grassmann structures. *J. math. Psychol.* **12**, 283–303.

Krantz D. H. (1975b) Color measurement and color theory: II. Opponent-colors theory. *J. math. Psychol.* **12**, 304–327.

Krauskopf J. (1974) Interaction of chromatic mechanisms in detection. *Mod. Probl. Ophthal.* **13**, 92–97.

Larimer J. L., Krantz D. H. and Cicerone C. (1975) Opponent process additivity. II. Yellow/blue equilibria and nonlinear models. *Vision Res.* **15**, 723–731.

Marc R. E. and Sperling H. G. (1977) Chromatic organization of primate cones. *Science* **196**, 454–456.

Miller S. S. (1972) Psychophysical estimates of visual pigment densities in red-green dichromats. *J. Physiol.* **223**, 89–107.

Mollon J. D. and Krauskopf J. (1973) Reaction time as a measure of the temporal response properties of individual color mechanisms. *Vision Res.* **13**, 27–40.

Mollon J. D. and Polden P. G. (1975) Colour illusion and evidence for interaction between colour mechanisms. *Nature* **258**, 421–422.

Mollon J. D. and Polden P. G. (1976) Absence of transient tritanopia after adaptation to very intense yellow light. *Nature* **259**, 570–572.

Mollon J. D. and Polden P. G. (1977a) Saturation of a retinal cone mechanism. *Nature* **265**, 243–246.

Mollon J. D. and Polden P. G. (1977b) An anomaly in the response of the eye to light of short wavelengths. *Phil. Trans. R. Soc. B* **278**, 207–240.

Mollon J. D. and Pugh E. N. Jr (1979) The anomalies of the blue-sensitive mechanism of the eye. In preparation.

Monasterio F. M. de, Gouras P. and Tolhurst D. J. (1975a) Trichromatic colour opponency in ganglion cells of the rhesus monkey retina. *J. Physiol.* **251**, 197–216.

Monasterio F. M. de, Gouras P. and Tolhurst D. J. (1975b) Concealed colour opponency in ganglion cells of the rhesus monkey retina. *J. Physiol.* **251**, 217–229.

Norman R. A. and Werblin F. S. (1973) Control of retinal sensitivity: I. Light and dark adaptation of vertebrate rods and cones. *J. gen. Physiol.* **63**, 37–61.

Polden P. G. and Mollon J. D. (1979) Negative masking of blue flashes. In preparation.

Pugh E. N. Jr (1976) The nature of the $\Pi_1$ colour mechanism of W. S. Stiles. *J. Physiol.* **257**, 713–747.

Pugh E. N. Jr and Sigel C. (1978) Evaluation of the candidacy of the $\Pi$-mechanisms of Stiles for color matching fundamentals. *Vision Res.* **18**, 317–330.

Rushton W. A. H. (1958) Kinetics of cone pigments measured objectivity on the living human fovea. *Ann. N.Y. Acad. Sci.* **74**, 291–304.

Stiles W. S. (1939) The directional sensitivity of the retina and the spectral sensitivities of the rods and cones. *Proc. R. Soc. B* **127**, 64–105.

Stiles W. S. (1946a) 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. (1946b) A modified Helmholtz line-element in brightness-color space. *Proc. Phys. Soc., Lond.* **58**, 41–65.

Stiles W. S. (1949a) Increment thresholds and the mechanisms of colour vision. *Doc. Ophthalmol.* **3**, 138–163.

Stiles W. S. (1949b) Investigations of the scotopic and trichromatic mechanisms of vision by the two-colour threshold technique. *Rev. Opt.* **28**, 215–237.

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.* **45**, 100–114.

Stiles W. S. (1967) Mechanism concepts in colour theory. Third Newton Lecture *J. Colour Group* No. 11. 106–123.

Stiles W. S. (1978) *Mechanisms of Color Vision.* Academic Press, London.

Stiles W. S. and Burch J. M. (1958) N.P.L. colour-matching investigation. Final report. *Optica Acta* **6**, 1–26.

Valeton J. M. and van Norren D. (1979) Transient tritanopia at the level of the ERG *b*-wave. *Vision Res.* In press.

Vos J. J. and Walraven P. L. (1971) On the derivation of the foveal receptor primaries. *Vision Res.* **11**, 799–818.

## APPENDIX I

### A. *Notation*

The following is a list of symbols used in the text, with the intended interpretation. The implicit context is that of the two-color threshold experiment.

$\lambda$ = wavelength [nm] of a monochromatic test flash

$\mu$ = wavelength [nm] of a monochromatic background or field

$U_\lambda$ = intensity [quanta·deg$^{-2}$·sec$^{-1}$] of the test flash at threshold

$W_\mu$ = intensity [quanta·deg$^{-2}$·sec$^{-1}$] of an adapting field

$A = A(\mu)$ = spectral irradiance distribution of a field composed of several too many monochromatic components

$\pi_{j\lambda}$ = test sensitivity [quanta·deg$^{-2}$·sec$^{-1}$]$^{-1}$ of the j$^{th}$ branch (Stiles's nomenclature) of the two-color threshold; the reciprocal quantum flux intensity at the absolute threshold of the j$^{th}$ branch

$\Pi_{j\mu}$ = field sensitivity [quanta·deg$^{-2}$·sec$^{-1}$]$^{-1}$ of the j$^{th}$ branch of the two-color threshold; the reciprocal field intensity required to produce a 10-fold loss in sensitivity in the j$^{th}$ branch

$\alpha_\lambda$, $\alpha_\mu$ = normalized absorption spectrum of the short-wavelength sensitive cones ($\lambda_{max} \simeq 435$ nm)

$\beta_\lambda$, $\beta_\mu$ = normalized absorption spectrum of the middle-wavelength sensitive cones ($\lambda_{max} \simeq 535$ nm)

$\gamma_\lambda$, $\gamma_\mu$ = normalized absorption spectrum of the long-wavelength sensitive cones ($\lambda_{max} \simeq 570$ nm)

$\alpha(A) = \int A(\mu)\alpha_\mu d_\mu$ = quantum flux density absorbed by $\alpha$-cones from the field $A$ of spectral irradiance $A(\mu)$

$\beta(A) = \int A(\mu)\beta_\mu d_\mu$ = quantum flux density absorbed by $\beta$-cones

$\gamma(A) = \int A(\mu)\gamma_\mu d_\mu$ = quantum flux density absorbed by $\gamma$-cones

$x$ = dimensionless variable, product of a field sensitivity and a quantum flux density

$\zeta(x)$ = the standard Stiles increment threshold function (Wyszecki and Stiles, 1967, p. 578) in linear coordinates

$g(A)$ = monotone decreasing steady-state "gain" function of the $\Pi_1$ pathway: see Eqns 1 and 2

$K_0$ = field sensitivity of first site, see Eqn 1

$K_1, K_2, K_3$ = "coupling coefficients" of $\alpha$-, $\beta$-, and $\gamma$-cones to second site with units of field sensitivity

$W_0$ = half-bleaching constant [quanta absorbed·deg$^{-2}$·sec$^{-1}$] (low density approximation; see Appendix III)

### B. *Definitions*

cone = a transfer function whose input is the spectrally integrated quantum flux density of one of the three absorption spectra $\alpha_\mu$, $\beta_\mu$, $\gamma_\mu$; the output is, at steady state, a monotone increasing or decreasing function of the total quantum catch

pathway = a concatenated sequence of transfer functions, the first of which is the transfer function of a single class of *cone*; each stage is imagined to behave as a low pass filter, stages past the first stage may receive input signals from more than one class of *cone*; the final stage is a *detector*, that gives a positive or negative response to a perturbation input to the pathway (see Fig. 9A,C)

site of adaptation = in a pathway, a transfer function whose gain may be altered by its own activity or the action of other events signalled to it; a perturbation input to the pathway probes the composite gain of all the sites of adaptation in the pathway

## APPENDIX II

### *Field Action Spectrum of $\Pi_1$*

The model considered here represents the steady-state gain of the two-stage system in the presence of a field $A = A(\mu)$ of arbitrary spectral composition by

$$g(A) = \zeta_1[K_0\alpha(A)]\cdot\zeta_2[|K_1\alpha(A)^n - K_2\beta(A)^n - K_3\gamma(A)^n|^{1/n}] \quad (1)$$

where $\zeta_1$ and $\zeta_2$ are the standard Stiles increment threshold function in linear coordinates (Wyszecki and Stiles, 1967, p. 578),

$$\alpha(A) = \int_0^\infty \alpha(\mu)A(\mu)\,d\mu, \quad \text{etc.}$$

represent the total quantum catches of the three receptor classes ($\alpha$, the short-wavelength sensitive receptors, $\beta$ and $\gamma$ the middle- and long-wavelength sensitive receptors, respectively), and where threshold elevation is given by $U_\lambda \pi_{1\lambda} = 1/g(A)$.

Here we have evaluated Eqn 1 by finding for the two observers of Pugh (1976), and for Stiles's average observer (Wyszecki and Stiles, 1967) the parameters $K_0$, $K_1$, $K_2$, $K_3$, $n$ that minimize the error function

$$E(K_0, K_1, K_2, K_3, n) = \sum_\mu \sum_x \{\log[\zeta(x)] - \log[g(W_\mu \oplus W_{\mu A})]\}^2$$

where $x = \Pi_{1\mu}W_\mu$, $W_\mu$ is the field intensity (in quanta$\cdot$deg$^{-2}\cdot$sec$^{-1}$), $\Pi_{1\mu}$ is the observer's field sensitivity at $\mu$, $W_{\mu A}$ represents the auxiliary field, $\oplus$ superposition of fields, and $g(\cdot)$ the function in Eqn 1. The wavenumbers considered were 24,500–16,000 cm$^{-1}$ inclusive in steps of 500 cm$^{-1}$; log $x$ was varied from $-2.8$ to $-0.8$ inclusive in steps of 0.2 log units. Thus, $18 \times 11 = 198$ squared deviations were calculated per function computation. In the $\frac{1}{2}$ log unit range of the increment threshold considered, the term $[K_1\alpha(A)]^n$ is poorly determined. We have thus constrained $K_1$ to satisfy, at $\mu = 500$ nm,

$$0 = (K_1\alpha_\mu)^n - (K_2\beta_\mu)^n - (K_3\gamma_\mu)^n$$

a constraint which yields important consistency of the model with increment threshold results outside the range considered in the fitting.

Table 1 shows the results of the fitting procedure, giving the parameters that yielded the best fits. The cone action spectra used were Stiles' $\Pi_3$, $\Pi_4$ and $\Pi_5$. In view of Pugh and Sigel's (1978) results there seems little reason to prefer any other set of fundamentals. We have also done the fitting with Vos and Walraven's (1971) middle- and long-wavelength fundamentals in place of $\Pi_4$ and $\Pi_5$; the goodness of fit is about the same, though the parameters differ somewhat.

There are several reasons for using in the error function the idealized increment threshold functions $\zeta(x) = \zeta(\Pi_\mu W_\mu)$ rather than fitting the model directly to available increment threshold curves. Firstly, the actual increment threshold curves are not available for Stiles's average observer (see Stiles, 1978, for details of the procedure used to obtain the $\Pi_1$ field sensitivity); thus, adoption of some such procedure was necessary at least to evaluate the model's performance with respect to the average observer. Secondly, Pugh (1976) found that between-day variability of $\Pi_1$ field sensitivity was not insubstantial, particularly in the long-wavelength spectral region. Thus, pooling increment threshold data across days is statistically less defensible than finding a mean field sensitivity by averaging the estimates obtained from individual increment threshold curves. Thirdly, the available evidence (Stiles, 1939, 1953; Pugh, 1976; see also Fig. 5, this paper) supports the conclusion that the $\Pi_1$ increment threshold curve within $\frac{1}{2}$ log unit of absolute threshold at every field wavelength is quite well approximated by the standard shape. The present fitting procedure is thus a logical and computationally efficient first-order method for evaluating the model's performance with respect to all available $\Pi_1$ field sensitivity results. The presentation of the curves generated by the model in Figs 5 and 7 with the standard shape provides a reasonable visual test of the model's ability to capture the observed approximate shape-invariance of the $\Pi_1$ increment threshold curves.

A difficulty with a previous analysis of a formulation similar to Eqn 1 (Pugh, 1976, Eqn 10) can now be understood as due to a failure to consider explicitly the contribution of the long-wavelength auxiliary fields. Even when such fields do not themselves elevate the $\Pi_1$ threshold, they can act to cancel signals fed by the $\alpha$ or short-wavelength cones to the second site. The auxiliary field thus prevents adaptation at the second site from being produced by the $\alpha$-cone signals under $\Pi_1$ isolation.

## APPENDIX III

### *Saturation of Steady-State Signals to the Second Site Due to Pigment Bleaching*

If we assume that cone signals in the steady state are functions of the rate at which photons are actually absorbed, then it is necessary at high intensities to allow for the bleaching of the photopigment, and the purpose of this appendix is to show formally how we have corrected the hypothetical signals to the second site to make them depend on the actual rate of absorption rather than on the intensity of the field. The principal equation involved here is the monomolecular kinetic equation for cone pigment of Rushton (1958). Alpern, Rushton and Torii (1970, p. 471) used this equation to account for the failure of cones to exhibit the increment threshold saturation observed by Aguilar and Stiles (1954) in rods. Output "saturation" of any steady-state cone signals to fields of increasing intensity is a consequence of the same bleaching equation. However, the concept of steady-state output "saturation" is completely distinct from increment threshold "saturation". We have included this appendix to clarify the distinction between these two concepts, and to show formally how Rushton's (1958) equation was incorporated into our description of signals to the second site. For simplicity we discuss first the case of pigment present in low density.

*Pigment in low density*

The differential equation describing bleaching and regeneration in the presence of a monochromatic field of wavelength $\mu$ and intensity $W_\mu$ is

$$\frac{dp}{dt} = -\epsilon_\mu \gamma W_\mu p + \frac{(1-p)}{t_0}$$

where $p$ is the fraction of pigment present and $\epsilon_\mu$ the extinction coefficient; $\epsilon_\mu\gamma$, the photosensitivity at wavelength $\mu$, is given in the inverse units of the flux density $W_\mu$ of the field. At equilibrium, the fraction of pigment present is

$$p_\infty = 1/(1 + \epsilon_\mu \gamma_\mu t_0 W_\mu).$$

The quantum flux absorbed, in the same units as $W_\mu$, in the steady state is

$$\begin{aligned} W_{abs} &= \epsilon_\mu W_\mu \cdot p_\infty \\ &= \epsilon_\mu W_\mu/(1 + \epsilon_\mu \gamma t_0 W_\mu) \\ &\rightarrow 1/\gamma t_0, \quad \text{as} \quad W_\mu \rightarrow \infty. \end{aligned}$$

Now we note that at intensity levels $W_\mu$ that produce insignificant bleaching $p_\infty \simeq 1$, and so $W_{abs} = \epsilon_\mu W_\mu$. In formulating the model we have implicitly made use of this proportionality, for we described the cone signals as functions of $(\epsilon_\mu/\epsilon_{max})W_\mu$ (i.e. $\alpha_\mu W_\mu$, $\beta_\mu W_\mu$, $\gamma_\mu W_\mu$—see Appendix I.) In other words, at low intensities we described the signals in terms of $(\epsilon_\mu/\epsilon_{max})W_\mu = W_{abs}/\epsilon_{max}$. To obtain a consistent definition of the steady-state signals to fields of such intensities that they give rise to non-trivial bleaching, we have substituted in Eqn 2 for $(\epsilon_\mu/\epsilon_{max})W_\mu$ the quantity

$$\begin{aligned} W_{abs}/\epsilon_{max} &= (\epsilon_\mu/\epsilon_{max}) \cdot W_\mu \cdot p_\infty \\ &= (\epsilon_\mu/\epsilon_{max}) \cdot W_\mu/(1 + \epsilon_\mu \gamma t_0 W_\mu). \end{aligned}$$

Clearly, at high intensities $W_{abs}/\epsilon_{max} \rightarrow W_0 = 1/\epsilon_{max}\gamma t_0$, where $W_0$ is the intensity of a light that produces a 50% bleach.

*Pigment present in density*

Primate cone pigments are now known to be present "in density" (Miller, 1972; King-Smith, 1973a, b; Bowmaker, Dartnall, Lythgoe and Mollon, 1978). It is thus worthwhile to inquire what effect this condition has on the equilibrium absorption rate. The steady-state bleaching equation involved has been discussed by Alpern and Pugh (1974, Eqn 7). It is not difficult to show that as the intensity $W_\mu$ of a monochromatic field increases the absorption rate (normalized by the dark-adapted absorption coefficient at the $\lambda_{max}$) satisfies, as $W_\mu \rightarrow \infty$,

$$\frac{W_{abs}}{1 - 10^{-D_{max}}} = \frac{\ln(10) \cdot D_{max}}{1 - 10^{-D_{max}}} \cdot W_0$$

where $W_0$ is the half-bleaching constant for the same pigment in dilute solution, and $D_{max}$ is the $\lambda_{max}$ optical density. Thus for $D_{max} = 0.55$ the term approaches 1.76 $W_0$. The principal difficulty of applying this analysis is that the kinetic descriptions of human *cone* pigments to date have all assumed low optical density. In effect, then, the correction factor for density has been absorbed in the cone pigment kinetic description.