Vision Res. Vol. 30, No. 11, pp. 1539–1553, 1990
Printed in Great Britain.
0042-6989/90 $3.00 + 0.00

DO DEVIATIONS FROM RADIANCE-INVARIANCE OF METAMERIC MATCHES CONTRADICT THE THREE PIGMENT THEORY OF FOVEAL TRICHROMACY?

MATHEW ALPERN, HUAZHONG ZHANG and JUN NOJI*
University of Michigan, Ann Arbor, MI 48109, U.S.A.

(*Received* 10 *August* 1989; *in revised form* 18 *January* 1990)

Abstract—According to the pigment theory of matching, metameric matches result from the equation of the rates of photoisomerizations for each of the three classes of cone pigments excited by the two matched fields. If true, matches are radiance-invariant and additive. Tests of the theory in this paper show small but ubiquitous failures in radiance-invariance due to systematic rather than random errors in matching. A choice between two possible explanations for these systematic errors favors the view that in subjects who deliberately or intuitively search for the middle of the matching range, the errors are due to an asymmetry in the Weber fraction for color (Trezona) at low (but not high) levels of retinal illuminance.

Metameric matches Isomeric matches Grassmann's laws of scalar multiplication Grassmann's laws of additivity Symmetric color matches Pigment theory of matching Foveal trichromacy

INTRODUCTION

The pigment theory of matching attributes foveal trichromatic color matches to the equation on the two sides of the colorimeter field of the rates of quanta absorbed by each of the three visual pigments in three corresponding photoreceptor species. If correct, matches must be radiance-invariant and additive (Grassmann's laws of scalar multiplication and additivity; Krantz, 1975; Pugh, 1988).

The theory is not generally valid: e.g. physically identical stimuli fail to match in viewing conditions influenced by spatial or temporal contrast effects. But it is believed true for matches visually equivalent by strict substitution (i.e. strictly symmetrical matches) or even those with quasi-symmetry† (Wyszecki & Stiles, 1982, pp. 278–285), despite evidence (Blottiau, 1947; Trezona, 1953, 1954; Stiles, 1955; Crawford, 1964; Wyszecki & Stiles, 1982, pp. 379–392; Zaidi, 1986) that it sometimes fails.

A variety of *ad hoc* hypotheses maintain the viability of the Pigment Theory in the face of these discrepancies. They include: (i) intrusion of rods (Alpern & Tamaki, 1983); (ii) nonlinearity in the responses of short-wave sensitive cones (Zaidi, 1986); and (iii) matching imprecision (Trezona, 1954; Stiles, 1955; Brindley, 1960). Rod intrusion can be dismissed if the colorimeter fields are viewed exclusively by the rod-free fovea, but the usual assumption that this is assured by central fixation of a small (≤ 2 deg in diameter) field is untenable (Alpern & Tamaki, 1983; Ahnelt, Kolb & Pflug, 1987). The theory must hold if the match equates quanta absorbances in three species of cones irrespective of whether the stimulus–response curve is *linear* or *nonlinear*, but a mismatch might easily be camouflaged as a match if the responses to it were on the saturated limb of a nonlinear curve. This approach to Zaidi's results merits further attention but there are other possibilities. Too little is known, for example, about the kinetics of bleaching and regeneration of cyanolabe, the short-wavelength-absorbing cone visual pigment. So we remain uncertain of the range of radiances over which its absorption spectrum is vulnerable to changes with radiance due to self-screening (Brindley, 1960; Alpern, 1979).

*Present address: Department of Ophthalmology, The Jikei University School of Medicine, Tokyo 105, Japan.

†In the elegant paper making this explicit Rushton (1972) drew a distinction between color *matching* to which the theory applied and color *appearance* to which it did not. This has the difficulty of implying a different meaning to the term visual *match* than that commonly employed (Wyszecki & Stiles, 1982, p. 278), i.e. the process whereby "... a determination is made of two physical stimuli that in some sense, produce the same response." This paper follows the latter terminology because a visual match must be determined on how its constituents appear if it is to be determined at all!

We here study the part of the spectrum in which short-wave-sensitive cones are not substantially involved in color matching. This paper tests Grassmann's laws of scalar multiplication and (to a lesser extent) additivity evaluating these alternative *ad hoc* hypotheses dismissing their failure.

METHOD

Maximum saturation centrally-fixated trichromatic color matches were obtained with apparatus and procedures already described (Alpern, Bastian, Pugh & Gras, 1976; Alpern, Kitahara & Krantz, 1983a; Alpern, Kitahara & Tamaki, 1983b; Alpern, Kitahara & Fielder, 1987). The task was to match the monochromatic test ($\lambda = 577.3$ nm, i.e. 17,322 cm^{-1} wavenumber) plus a desaturating primary, to a mixture of the other two primaries, by adjusting the radiances of the primaries alone until the two fields looked identical. The standard set of instrument (and reference) primaries: 15,500 cm^{-1} (645.2 nm), 19,000 cm^{-1} (526.3 nm), 22,500 cm^{-1} (444.4 nm) and WDW normalizations at 17,250 (579.7 nm) and 20,500 cm^{-1} (487.8 nm) were used.

Six observers were studied in preliminary matching experiments (at several test wavenumbers in the red–green spectral range), but only three endured enough experimental repetitions to allow radiance invariance to be put to a satisfactory statistical analysis. A sectored disc rotating at a high frequency in the final common path of the colorimeter attenuated radiance level of all match constituents equally (Alpern et al., 1983a).

To exclude the possibility that the effects were due to rods the test was often centered on a 5 deg (22,789 cm^{-1}, 438.8 nm) violet background of 3.6–3.9 log scot. td, depending on the subject. This insured that any rods in the image of the field were saturated. To exclude the contribution of short-wave sensitive cones to radiance-invariance failures, matches were repeated with the green primary changed to one of $\lambda_{max} = 550.3$ nm (i.e. 18,172 cm^{-1}). Observers then could match the test with a bichromatic mixture of the long- and middle-wave primaries alone; no desaturating short-wave primary was required. If short-wave sensitive cone nonlinearity is the only source for the failure in Grassmann's laws, radiance-invariance will apply.

These experiments extended over a period of 3 yr while the initially crude matching gradually refined. In the earliest phase, observers matched five times at the highest level starting from a random mismatch each time. The process was repeated at each successively lower radiance, in the same session. The first modification of this routine was the immediate average of the settings of the highest radiance matches before proceeding to the next highest. Observers confronted the average match of the first set and decided if it was still a perfect match at the lower radiance. If not, the procedure was repeated at this level. But if the match held, the radiance was immediately reduced to the third level where the process was repeated. We continued in this way until all radiance levels were tested both without and with the background, in one session.

HZ, one of three subjects in this phase of evolution of the procedure was able to finish 15 daily sessions. In his case alone, the average results of the five settings at each level on a given day were used in the analysis of variance (one way: $d.f._1 = 6$, $d.f._2 = 84$). These results allowed a significant level of rejection of the null hypothesis and are included below (in Fig. 1 and Appendix Table A1). All the other data obtained by these first two methods treating each of the five settings (with or without the background) as a separate run when analyzed appropriately based on five daily means at each radiance level were not significant because of large daily variation in the pattern of scalar multiplication violations (though inappropriate analysis using 25 "runs" gave spurious significance). Factors producing between session variability interacted with radiance showing different violations of scalar multiplication from one session to another. Hence these data were not included in this paper, and further modification of the protocol was introduced.

Most of the results were obtained in the third phase of evolution of the matching process. After one match at the highest level, the first step down was introduced and, following adaptation to the new level, the subject was asked if the match still held. If so, the next step down was introduced and the process repeated. If not, he readjusted the settings for a match before the next step down. The routine continued until the entire sequence of down steps was finished. The run ended with repetition of the entire process on the background, 24 runs completed the set.

Statistical analyses are summarized in the Appendix Tables A1 (metameric matches) and A2 (isomeric matches). Though mean results only are plotted in Figs 1 and 2, the statistical analysis depended on the number of runs completed per session. For example, in Table A1 the first analysis (HZ standard primaries) was the result of 5 runs/session but the average of the five was used as the single result of that session. There were 15 sessions. FL (standard primaries) was only able to complete one run during the first 4 sessions but completed the rest of the set in 2 runs/session. Only the latter 20 runs are included in the analysis in Table A1, though the means of all 24 are plotted in Fig. 1. DW completed 2 runs/session with both metamere and isomere matches; HZ completed 4 runs/session (6 sessions). FL also completed 4 runs/session (6 sessions) for the metameric matches with the more reddish middle-wave primary but matched isomeres with the earlier metameric matches using the standard primaries so that for the isomeric matches only 20 runs (2 runs/session) were included in the analysis (Table A2).

The two colorimeter fields were presented to the same patch of retina (a 1 deg circle) in 1 sec successive exposures for as long as needed. This (alternate presentation) mode was chosen for two reasons: (i) the matches were sufficiently difficult that the completion of any one required enough time that the assumption of quasi-symmetry was less tenuous than the case for its alternative; (ii) simultaneously exposed colorimeter fields have the difficulty emphasized by McCree (1960) that staring at the field without interruption causes mismatches in the display to fade. This makes what we call a pigment match (i.e. *the settings of the primaries the result of which fulfills the predictions of the pigment theory of matching*) more difficult with the two fields simultaneously displayed. Every effort was made to optimize pigment matches.

Except as noted, data were analyzed by two-way analysis of variance (mixed model–sessions and radiance—with nested random effects) (Winer, 1971).

RESULTS

(A) Scalar multiplication

Figure 1 shows mean matches with standard primaries. Each graph is a plot of WDW chromaticity (u—the long-wave coordinate at the top, v—the middle-wave coordinate in the middle, and w—the short-wave coordinate, below) as a function of test radiance. Symbols represent different subjects, on the left without, on the right with the background. Included as triangles among the former are Stiles' 1955 data matching 581.4 nm test with a 2 deg field.* (Though means of two subjects, the results are replotted as in the original as if each point were one match; no statistical inference can be drawn from them.)

Summaries of all other data in this and the following figure are shown in Appendix Tables A1 and A2 where it may be noted from the designated significance levels of the F-ratios (error term: radiance × sessions) that without the background the results differ significantly from expectations of scalar multiplication (i.e. a straight horizontal line) for each subject.†

Two points are notable about these small deviations from the scalar multiplication description of data in Fig. 1. The first is how small they are. Stiles' data are relevant in the context of their size. Stiles dismissed the deviations from horizontal he found because of the small size: "... except for a slight drop in the positive red ...", with which however he fails further to deal. Yet this slight drop in positive red is a larger deviation from the expectation of scalar multiplication than any we found. The second is that, despite this small size, all the results are statistically significant. Thus we cannot dismiss these deviations as due to purely random measurement imprecision. The statistical reliability of these deviations from a null effect also constitute statistical evidence that there are *not* large deviations from the null effect.‡ Whatever the cause, the generality of the *presence* of this small effect (though its *direction* with decrease in test light intensity seems random among observers) implies that there is a systematic deviation from scalar multiplication for each observer.

*We took Stiles' data directly from the graph in his paper and plotted them as accurately as possible. They are not quite correct (the sum of the three coordinates is sometimes ≠ 1.0, an impossibility by definition), but we do not know how to improve the plotted points estimated by reading a magnified version of the original data assisted by a bit pad.

†Remarkable is that the estimates of variance for radiance × runs (w sessions) are not appreciably smaller than those for radiance × sessions. Thus improvements in the protocol in advancing to phase 3 from phase 2 eliminated any statistically detectable between-session variability in the pattern of radiance effects. This made it unnecessary to make the sharp distinction between sessions and runs as sources of variability we were obliged to introduce as a consequence of the results of phase 2.

‡We are indebted to Professor D. H. Krantz for pointing this out to us.

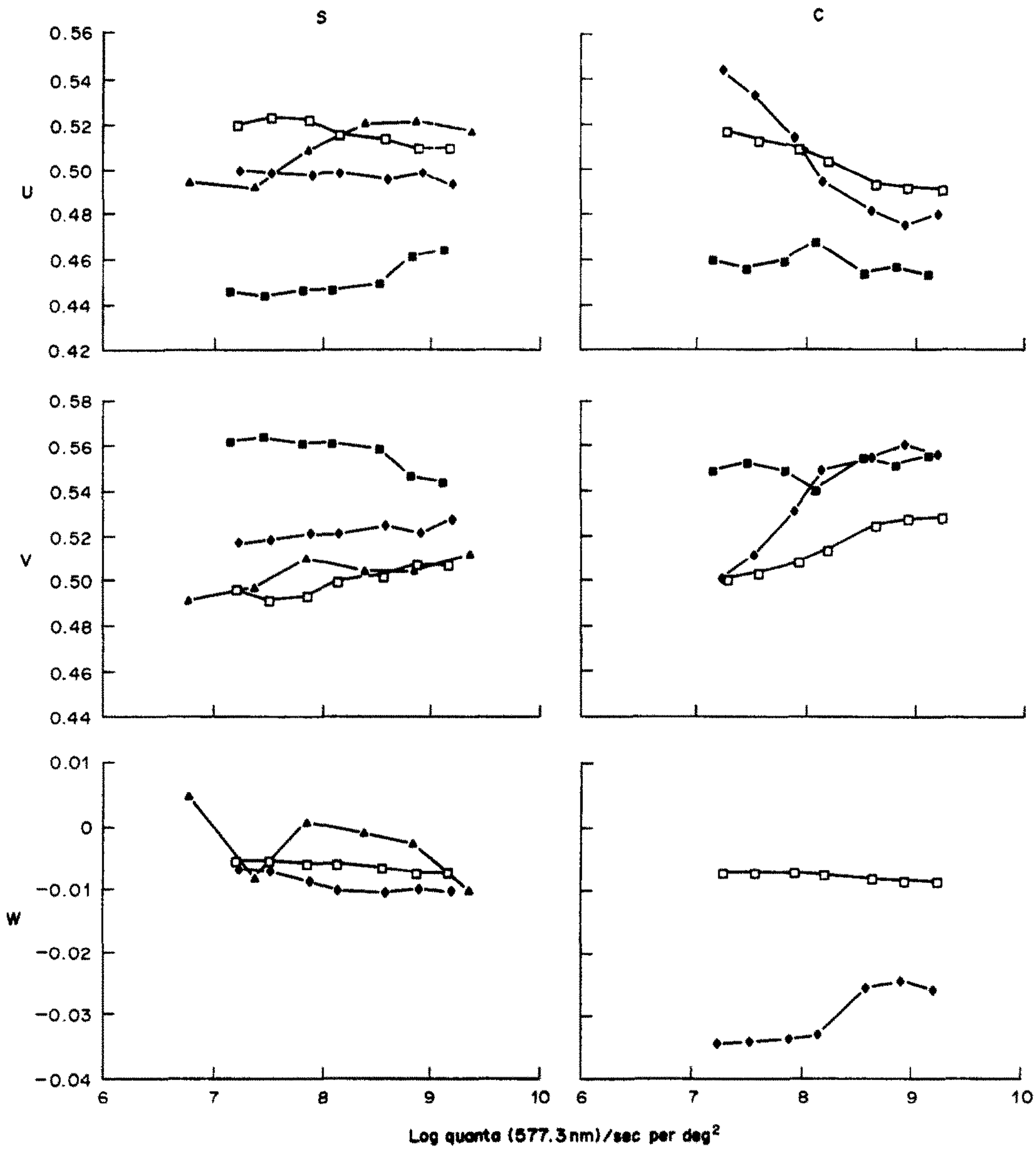

Fig. 1. Test of scalar multiplication of a yellow test light for the subjects of this study, together with analogous data (▲) from Stiles (1955). WDW chromaticities (top row: *u* or "red"; middle row: *v* or "green"; bottom row: *w* or "blue") each plotted as a function of the log of the test radiance on the left without, on the right with, the background (μ). Stiles tested 581.4 nm. The abscissas can be converted to trolands by deducting 6.1644 (for 577.3 nm), or 6.1863 (for 581.4 nm). The results for HZ are the means of 15 sessions each session containing five runs averaged and treated as a single run in the analysis. The results for FL and DW are means of 24 runs in the latter case 2 runs/session, the former were obtained in 1 run/session for the first 4 sessions and 2 runs/session for the final 20 runs. Stiles' data were taken from the graphs in his paper. There seem to be errors in plotting at least the final two points which do not add to unity as they are normalized to do. Such minor errors in the third decimal are not sufficient to dismiss the failure of scalar multiplication conceded in his text. Symbols: □, HZ; ◆, FL; ■, DW; *s* without, *c* with the rod saturating background. Instrument primaries equal to the reference (standard) primaries.

The accepted explanation for the (much larger) deviations with 10 deg fields is rod intrusion in the match. Can a similar explanation deal with the left-hand set of results in Fig. 1 (with the differences in the sizes of discrepancies due merely to the relatively small number of rods excited by the 1 deg field)? The results on the right in Fig. 1 show that for only one subject (DW) is such an explanation tenable. The others persisted in showing significant deviations even on rod-saturating backgrounds.

(B) Contribution of short-wave sensitive cones

It is not clear how a nonlinearity in short-wave sensitive cones (Zaidi, 1986) might deal with the failures of scalar multiplication evident in Fig. 1. Still all three primaries participated in the match (with the notable exception of DW who used no short-wavelength desaturant to match 577.3 nm with a mixture of the 19,000 cm^{-1} and the 15,500 cm^{-1} primaries), so the results do not exclude hypothetical short-wave cone nonlinearities as sources for the failures of scalar multiplication observed.

For this exclusion, we repeated these experiments on HZ and FL with the 18,172 cm^{-1} instead of 19,000 cm^{-1} green primary. Both then made satisfactory matches to the yellow test with no desaturating blue primary.

Diamonds and squares in Fig. 2 are the results. As before, the first column of graphs in Fig. 2 are without (s), those in the second are with (c), the background. They are analyzed below in Appendix Table A1. All subjects show radiance-invariance failure with no blue cone input. For DW the possibility that this was due to rods cannot be excluded, for it disappeared

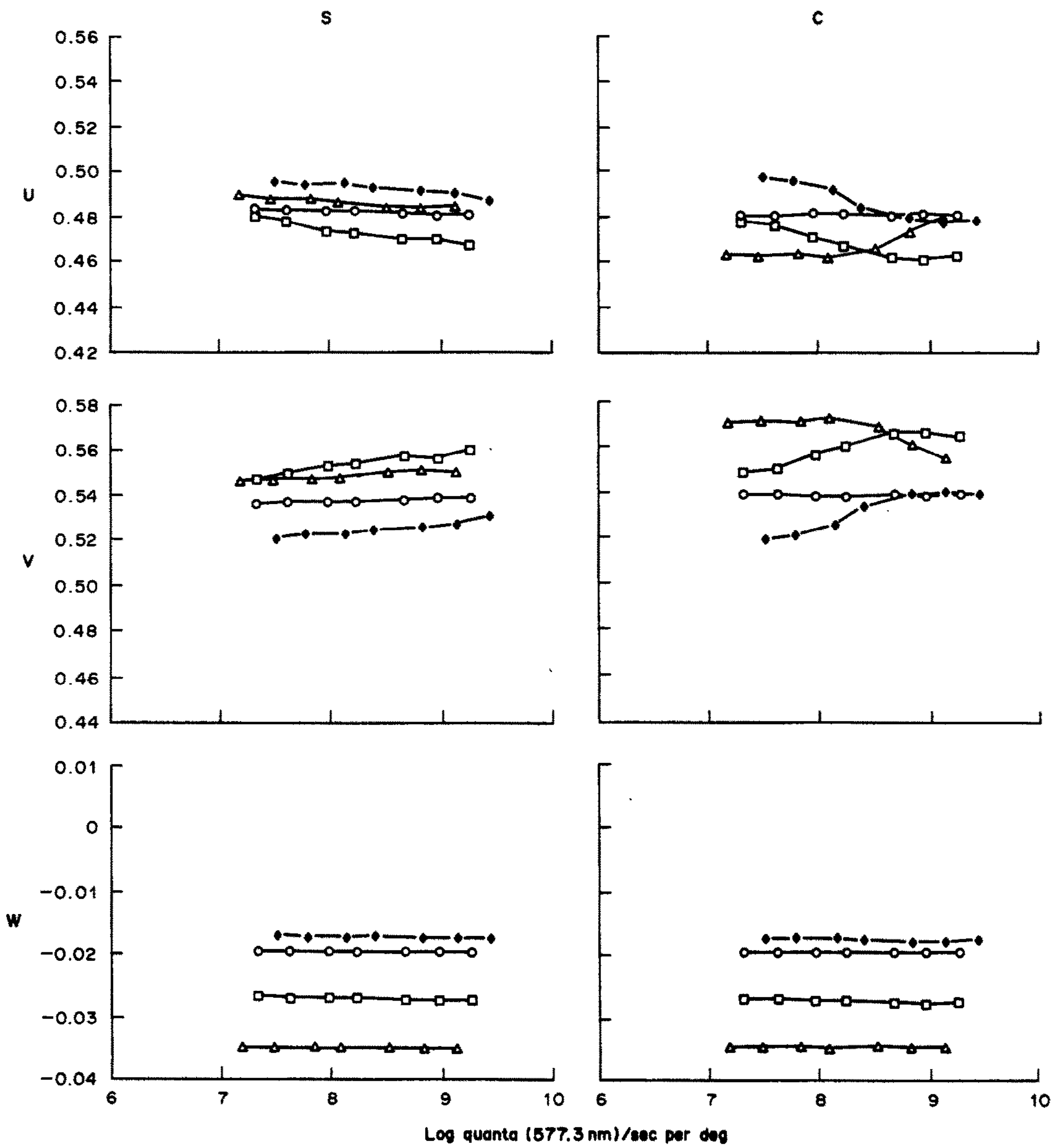

Fig. 2. WDW chromaticities measured with the green instrument primary changed to 550.3 nm, a wavelength sufficiently long that no desaturating blue primary was required in the match. Though measured with different instrument primaries the data have been transposed into the same *reference* primaries as those in Fig. 1 before the WDW normalization. Symbols and conventions otherwise as in the legend to Fig. 1. The results of matching apparent isomeres are shown as (○) for HZ and (△) for FL. Note that the latter set measured on the background deviates substantially (and significantly) in the opposite direction from comparable metameric matches.

on a rod saturating background; this is not the case for HZ and FL.

(C) Apparently isomeric matches

Consider an ideal observer with zero dead zone of indiscriminability matching according to the theory. There is only one unique match equating the rate of photon absorption for all three cone species. Grassman's laws must hold. Actual observers differ from the ideal in that each receptor species has a finite zone of indiscriminability (in which mismatches are indiscriminable from pigment matches). It is easier to grasp this complexity by considering the case when the match is independent of the detectors: e.g. when the match is a physical match (i.e. an *isomeric* match).

Subjects set the wavelength drum of the colorimeter double monochromator (± 1.0 nm HBW) to match the light in a colorimeter channel attenuated by an interference filter (± 4.5 nm HBW) $\lambda_{max} = 577.3$ nm. The radiance of the monochromator beam was also set if needed. Once the first match was achieved, the experiment proceeded with the usual routine, i.e. the smallest step down was introduced with the rotating sector and the subject judged whether the match held. If not, he corrected it with the wavelength alone if possible or (more rarely) with both wavelength and radiance adjustments. The next step down was introduced and so on until the entire range of steps was covered first without, then with the background. To compare *isomeric* matches with *metameric* matching data, wavelength in nanometers, was converted to WDW chromaticities by interpolation in a table (unique for each observer) of the mean metameric matches with the standard set of primaries at ~ 10 nm intervals throughout the visible spectrum.

Isomeric matches are shown in Fig. 2 as open triangles for FL, and open circles for HZ. The results of the analysis of variance of the isomeric matches are given in Appendix Table A2. Scalar multiplication held for DW's isomeric matches and for those of HZ on the background.* All other data in Fig. 2 show significant radiance-invariance failures in both metameric and isomeric matches. Those of the metameres were usually larger and for FL on the background, in the opposite direction from deviations uncovered by isomeres. The former is consistent with the hypothesis that the small but significant deviations from scalar multiplication deviate from a pigment match because real observers depart from the ideal with its unrealistic zero deadzone of indiscriminability. But if this is the explanation, how is it that these deviations are not random, as commonly supposed (Wyszecki & Stiles, 1982, p. 390; Zaidi, 1986), but systematic. We consider this next.

*We cannot believe that the extremely small deviations from scalar multiplication in HZ's settings of the w chromaticity with the background (though significant statististically) have any perceptual meaning.

PART II. HOW DOES RADIANCE-INVARIANCE SYSTEMATICALLY FAIL?

Two explanations for systematic failures of radiance-invariance of metameric matches, both subsets of the subliminal mismatch hypothesis, have been proposed: (1) nonlinearity of the stimulus–response function of cone mechanisms (Zaidi, 1986) and (2) asymmetry of the Weber fraction for color (Trezona, 1954).

The mismatch postulated in the first possibility is explicit on p. 1539 that it "... might easily be camouflaged as a match if the responses to it were on the saturated limb of a nonlinear curve." (The details are ours, not Zaidi's who is explicit only that the nonlinearity responsible for the failure in additivity he examined was specific for blue cones. Clearly, the data in Fig. 2 are not explained by a nonlinearity for blue cones.)

To understand the second possibility, consider the following experiment. Adjusting only a single primary in the colorimeter field (and holding brightness fixed) Trezona (1954) found that the size of the step from a perfect match to just "too blue" was smaller than the size of the step to the just "... too little blue". So the perfect match was not in the middle of the matching range. If the perfect match was a pigment match, then the latter would be asymmetrically distributed in the matching range. If observers either deliberately or intuitively searched for the midpoint of the matching range, then small mismatches could occur which lead to systematic violations of scalar multiplication; for Trezona found this only if the requisite blue in the match was small. Given the shape of the increment threshold versus radiance curves of Stiles' color mechanisms, it would not be surprising to find some (low) radiance level for which asymmetry of the Weber fraction for color is found.

A better understanding of these possibilities requires a way to specify a pigment match more precisely. Indeed Trezona's hypothesis leaves one skeptical of the possibility of achieving a pigment match by any psychophysical method. At least it is essential to verify radiance-invariance without permitting the subject to change the match at more than one radiance level. Can one ever make a valid match over a variety of levels with these constraints?

METHOD

The 4th phase in the search for an accurate method for making a pigment match was carried out on HZ, the most experienced and precise subject available. He adjusted the primaries to match the highest test radiance as before with no explicit instructions as to how this should be done other than the emphasis on an "exact match". Also, as before, the first step down was introduced immediately after a single match and the routine continued step by step downward with judgments at each step as long as the match remained exact with no change in its constituents. If the match no longer held, the primaries were readjusted to match as before. Instead of stepping down then to the next lower level, the routine was changed to retrace to the higher level to see if the new settings matched at the level where previously different settings were required. If so the experiment proceeded to the next (previously evaluated) step up, if not the primaries were set again. Resetting (as needed) of the primaries continued with matches retracing the steps first back to the starting point with "exact matches" at each step with no adjustment of primaries before receding in further down steps in search of an "exact match" at each of the seven test radiance levels with no adjustment of the primaries after the initial settings in the chain.

RESULTS

Initial trials were frustrating; no quantitative data were kept. No single match satisfied at all levels. Matches at the four highest levels were about the same, but they were distinctly different than those at lower levels. These first experiments ended after about an hour with headaches and feelings of helplessness over the inability to find a single match satisfactory for all levels. The 3rd and 4th attempts were undertaken 3 weeks later, on one day without $(s\mu)$, the next with $(c\mu)$, the background. Results were recorded; the pattern of inability to find a single match is shown in Fig. 3.

Each symbol in this (and the following) figure(s) are settings at a given test radiance, open, if made without background, solid, if made on the background. Three months later he tried again, first with the same lack of success but at the end of that fifth attempt HZ achieved radiance-invariance both without and with the background. A week later we came to the same conclusion after only a slight mismatch at the start of the session (Fig. 4).

Note in Fig. 4 that the match without the background for which scalar multiplication holds is not the final match *with* the background. Though scalar multiplication held for both sets of matches, at least one cannot be a pigment match. For additivity fails! The same result was noted in the previous week's session. A week later the test was repeated beginning with the settings of the primaries on the background at the value they had at the end of the run without the background. The transition time between the two sets was reduced to the amount of time (~ 1 min) needed to adapt to the background. Reaching the initial match for which scalar multiplication was valid without the background did not come as easily as in the preceding week. Once reached, the same match was valid both with and without the background. This was the first hint in 3 yr of search that HZ could set a metameric match so that both additivity and radiance-invariance hold.

More general is the 5th and final phase in the search for a valid way to test the pigment theory i.e. abandoning a test of radiance-invariance without a background *before* evaluating additivity. The additivity test was interleaved at one radiance level before studying the next. The protocol required adapting to the each set of conditions before any test could be undertaken and was more tedious than previous routines. It proved more difficult for HZ to manage. His first attempt to achieve it (Fig. 5) failed.

The next week he all but succeeded. The first match on the background was not right, but after readjusting it slightly, the next 12 were exact matches; only the final match, the lowest level of retinal illuminance without a background was not an exact match. Unfortunately the experiment was terminated rather than continued relentlessly. Finally, on the first trial a week later, the desired match was made, as Fig. 6 shows.

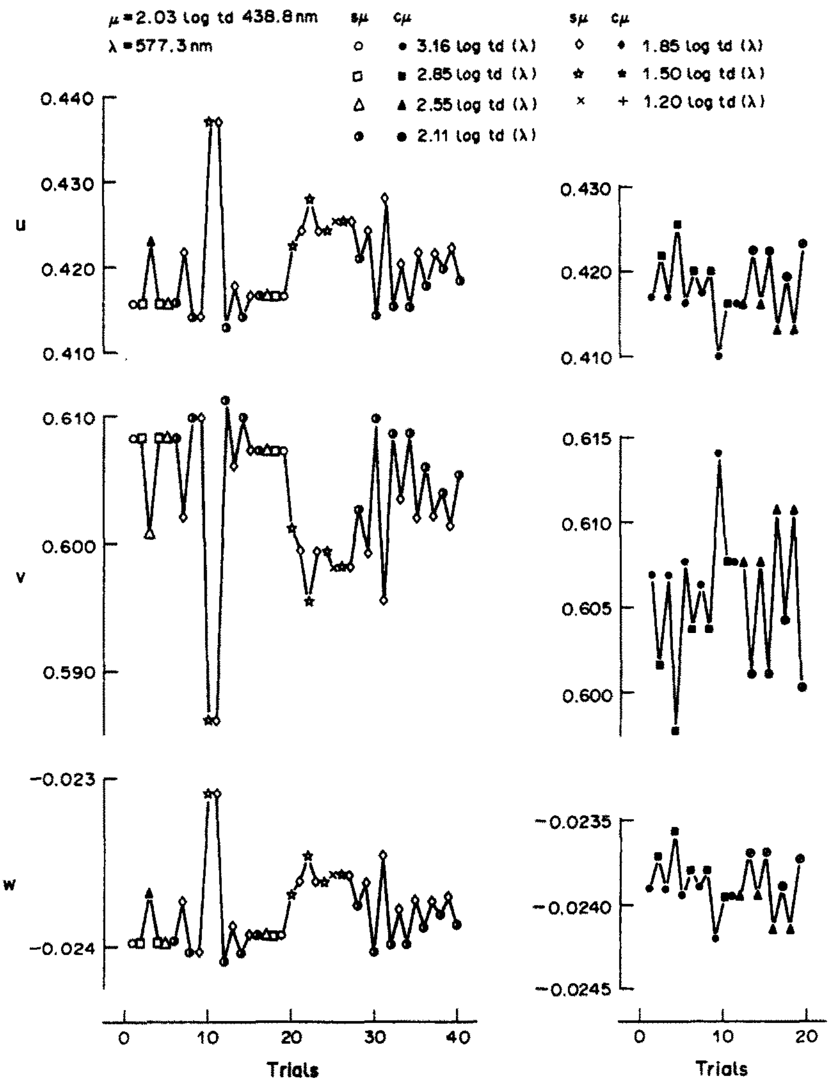

Fig. 3. Results of the 3rd (on the left) and 4th (on the right) attempt to find a pigment match with the protocol of Part II. Solid and open symbols, otherwise identical, show results at the same level with ($c\mu$), and without ($s\mu$) the background, respectively. Different sizes and forms of the symbols represent different retinal illuminances indicated above. After the initial settings only chromaticities differing from their respectively immediate left-hand neighbors resulted from adjustment of the primaries, the rest were obtained after inspecting the preceding match and confirmation that it still held. Data on the left were obtained on the date indicated, the right hand set on the next day. Instrument primaries are those used to obtain the data in Fig. 2.

The data in Fig. 6 prove that a psychophysical method can be realized with the operations we used to defined a pigment match. According to the theory, repetitions of these operations lead to the same value. But this expectation has never been confirmed; we took advantage of HZ's new skill to test this prediction. The result in Fig. 6 has been repeated on five occasions in 75 days. Only two of these, a pair obtained on the same day, yielded exactly the same match. The measurements, as for all the results in Figs 3–6 were made with the 550.3 nm primary, reducing the number of primaries from 3 to 2. The transformation expressing the data in the Stiles primaries naturally introduces a 3rd color matching function but for the present purpose this transformation was *not* used. The results were expressed as chromaticity in the instrument primary color space. Since the matches were dichromatic only a single chromaticity coordinate, say r, (the long-wave coordinate) need be calculated. The mean ± 1 standard deviation of the five matches was 0.6873 ± 0.0339.

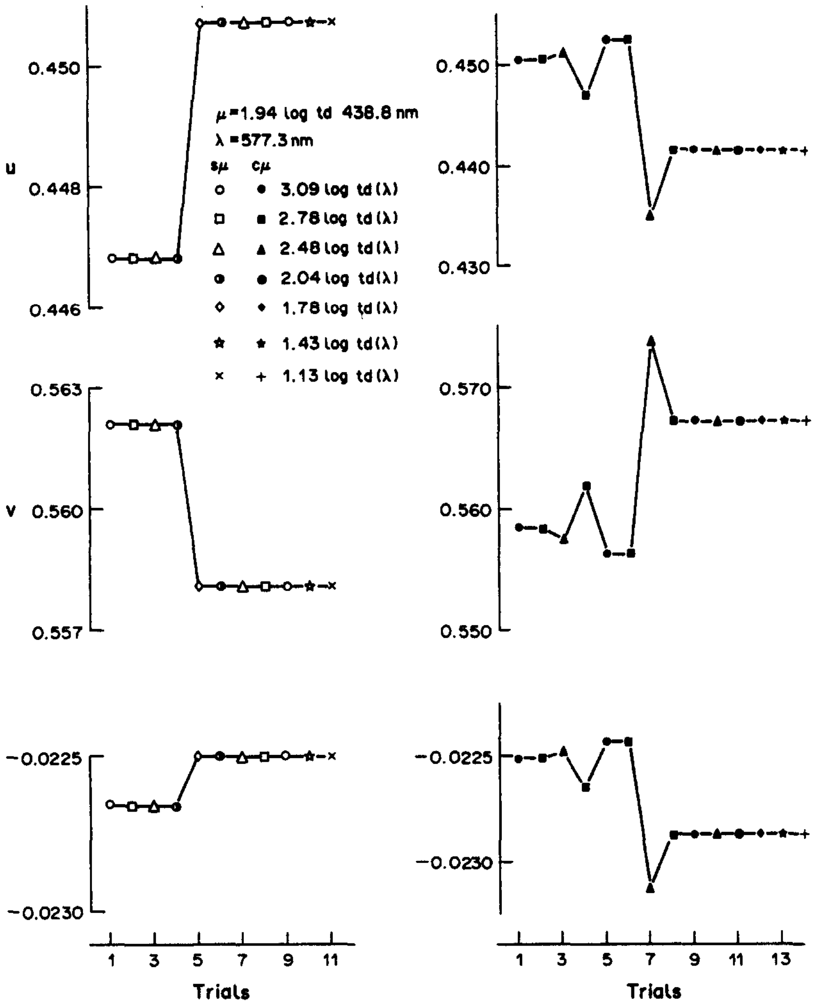

Fig. 4. Results of the 5th and 6th attempts to follow the protocol of Part II. The details are outlined in the legend to Fig. 3 except that HZ completed the runs with, and wihtout the background on the same day, with a rest interval between and a fresh initial match beginning the 2nd set. At the end of each run scalar multiplication held, but additivity failed! The match without the background wasn't identical to the match on the background.

DISCUSSION

Given the tiny size of the deviations found and their ubiquity, the ideal concluding experiment should be the confirmation that the match in Fig. 6 is a pigment match with the demonstration that the 2 half-fields of the colorimeter at the match are *indiscriminable* by 50% performance in a two alternative forced choice task as Cornsweet (1970) suggested for discriminating between trichromacy and tetrachromacy of color matches in the peripheral retina. Technical difficulties, hinted at below, have so far prevented us, or anyone else, from successfully completing such a demonstration and Trezona (1973, 1974) has made the discrimination which was Cornsweet's concern with the scalar multiplication test used here. But we have not yet despaired that these difficulties will eventually be mastered.

As this work progressed it was more and more evident that the idea of a single unique pigment match due exclusively to the equation of absorbances of three species of cone pigments excited by two colorimeter fields is too simple. It is no longer heuristic to suppose only three photoreceptors pairs are involved. In real eyes 1 deg fields excite (conservatively) over 1600 long-wave sensitive cones, perhaps an equal number of middle-wave cones and ~300 short-wave sensitive cones. Nor is it realistic to imagine the spectrum of prereceptor absorption across this 0.0665 mm^2 area is everywhere uniform. Nor is the eye immobile during the 1 sec exposure of the field; all these numbers are increased by eye movements. Changes in the

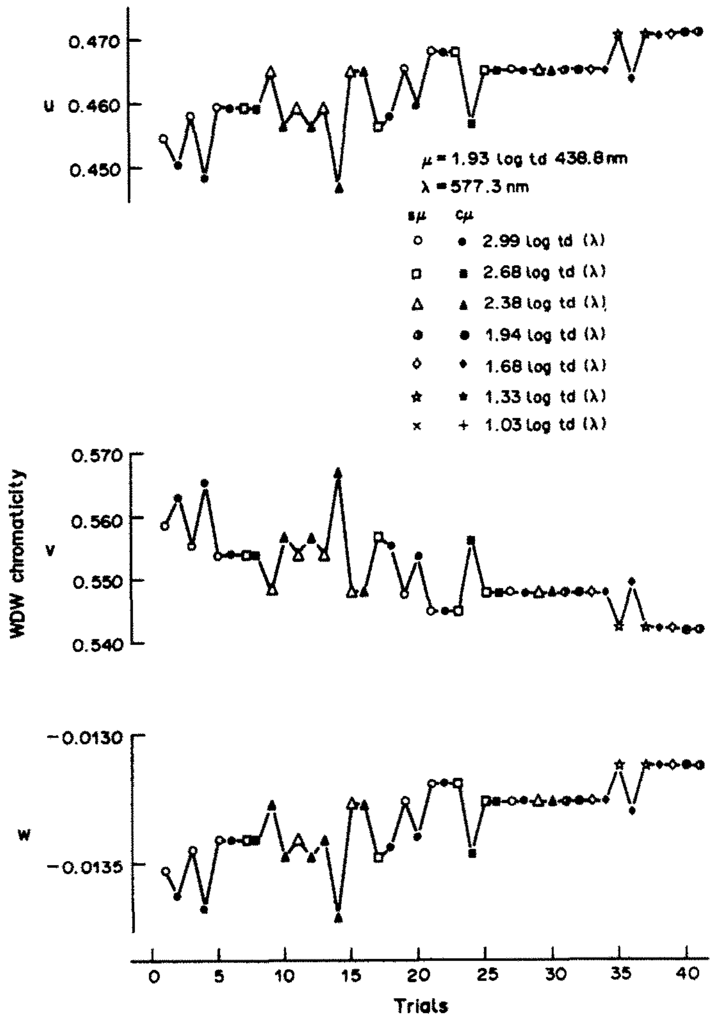

Fig. 5. First results with the final protocol. The procedure was the one used to obtain the results in Fig. 4 except that at every level of test retinal illuminance the match without the background was either followed or preceded by a match with the background at the same test intensity, care being taken to be fully adapted to the background before each match or judgment of a previous setting. No satisfactory match obeying additivity and radiance-invariance was found even after 40 trials.

angle of retinal incidence of the test inevitably follow these changes in fixation and they, in turn, result in changes in the action spectra of the three species of cones each small test beam component excites (Stiles, 1937; Enoch & Stiles, 1961; Brindley, 1953; Alpern et al., 1987; Alpern, 1986, 1989). Changes in these spectra are also due to small differences in outer segment length among receptors of the same species. Added to all of this are nonuniformities of the fields and inevitable imprecisions of calibrations. That the only identical matches in five repetitions of Fig. 6 were the pair relying on the same calibration suggests that calibration errors provide a lower limit to the distribution of putative pigment matches.

Considering the simple theory of pigment matching in view of these facts of real world colorimetry, we see what it needs in order to manage them is the hypothetical construct psychologists call *match criterion.*

Colorimeter fields are not uniform but blotchy. Part of the blotches come from nonuniformities in the colorimeter's optics and can be dealt with to some extent by attention to details in cleaning the optical surfaces of the apparatus; others are due to nonuniformities in the eye, prereceptor and receptors and are not so easily dismissed. In the end the decision about these blotches remains an usually unmentioned aspect of the match criterion; a change in the criterion, i.e. *a criterion shift*, is an important uncontrolled variable that even highly skilled subjects may find difficult to avoid. It is not surprising that small failures in the matching of physically identical lights are

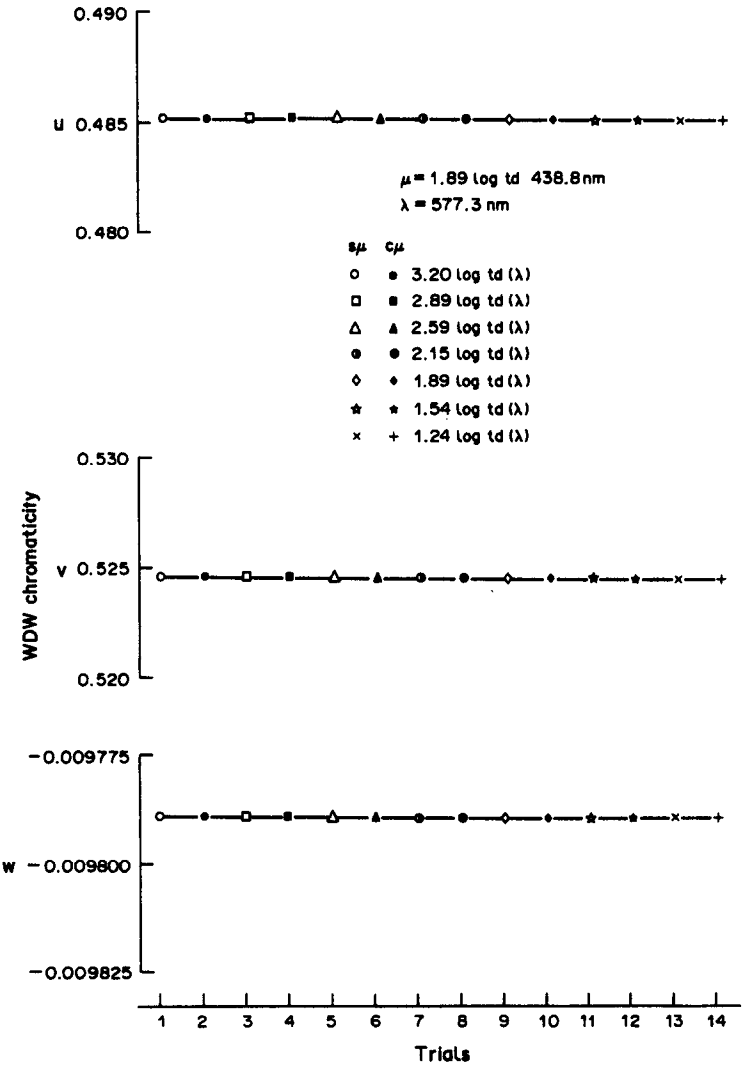

Fig. 6. Results of the 3rd attempt to follow the final protocol. Details of the procedure are given in the legend to Fig. 5. Additivity and scalar multiplication were found for the first adjustment of the primaries.

found in this (cf. the isomeric matches in Fig. 2 and Appendix Table A2) and nearly every other paper describing metameric matching in which radiances are measured.

So it is hardly unexpected that successive repetitions of a putative pigment match yield a distribution of matches, rather than the identical match. It is heuristic to return now to explanations for systematic failure of radiance-invariance mentioned above. The two different explanations are not, of course, mutually exclusive but if we make the simplifying assumption that, for a given observer one of the two will be more influential, that five repetitions of a putative pigment match are not identical, points to a simple way the more influential explanation may be identified. If a pigment match is variable then ordinary matches will also be variable. Under optimal conditions the two distributions should be about the same. The two hypothetical sources of radiance-invariance failure can only add to this variability of the pigment match. So according to either view as radiance changes a level will be reached where the distribution will broaden to include a distribution of subliminal mismatches responsible for the failure of radiance-invariance along with the pigment matches. The two hypotheses differ in the radiance level where this broader distribution is expected. In the nonlinear response hypothesis it is at the highest level where the curve is closest to the high radiance saturated limb. On the other hand, if observers match by searching for the middle of the uncertainty range in Trezona's hypothesis of the failure of symmetry of the Weber fraction for color, the broadest distribution of matches will occur at a low radiance level where two-color increment threshold vs

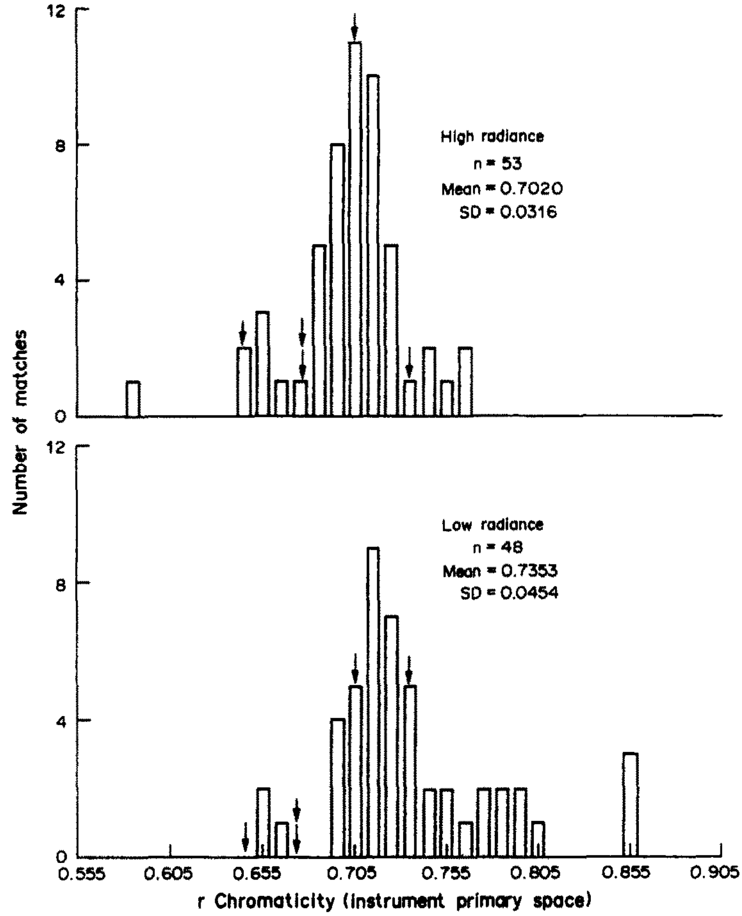

Fig. 7. Bar histograms: distributions of r the long-wave chromaticity in instrument primary color space of all (dichromatic) matches at the highest (top figure, $n = 53$) and lowest (bottom figure, $n = 48$) radiance levels by HZ. Arrows show the distribution of the five putative pigment matches made with experiment similar to those yielding the data in Fig. 6.

radiance curves shows remarkable differences in slope as radiance is changed.

To choose between these possibilities, the distribution of all HZ's ordinary matches are shown as bar histograms in Fig. 7 at the lowest (below), and highest (above), radiances. Also shown in both graphs by arrows is the distribution of the five hypothetical pigment matches.* This latter differs significantly from the distribution of ordinary matches at the lowest radiance (below, $t = 2.29$, 51 d.f., $0.02 < P < 0.05$) but not from that of ordinary matches at the highest radiance (above, $t = 0.99$, 56 d.f., $0.2 < P < 0.4$). (If the five points in the bar graph above obtained without a corresponding match at low radiance are dropped, the mean changes to 0.7067 ± 0.0254, but the inference about the difference between the two distributions above is unchanged, $t = 1.59$, 51 d.f., $0.1 < P < 0.2$.)

Thus the general trends shown in Fig. 7 follow the pattern expected of subjects searching for the midpoint of the uncertainty range in matching and Trezona's asymmetry of the Weber fraction for color at low radiances. According to the hypothesis attributing failures of radiance-invariance to mismatches on the saturated high radiance end of the response curve, ordinary matches should be more closely confined to the pigment match distribution once radiance is reduced to a level where the response curve is no longer on the saturated nonlinear limb. A pattern opposite to that seen in Fig. 7 may reasonably be expected according to this hypothesis. Given that so far the putative pigment matches have only been made by one-

*Despite the need for more repetitions of the result in Fig. 6 before a definitive discussion of differences between distributions of putative pigment and ordinary matches is realistic, HZ can no longer work on this problem. Hence the obligation to draw tentative inferences on the basis of available data, however necessary further experiments on others have become.

subject and the one order of magnitude differences in the sample size between presumed "pigment" and ordinary matches on *him*, this conclusion must remain tentative pending further studies of this kind.

However those experiments turn out, we infer that small mismatches from the expectations of the theory due to criterion shifts are responsible for the failures of Grassmann's laws found here. The evidence for this view is reasonably convincing as it relates to radiance-invariance. For additivity one must be more cautious. Only a few examinations of this property were undertaken.

Readers may wonder the extent the ubiquitous failures in scalar multiplication were due to imperfections in the apparatus, lack of subtle improvements making the fields more uniform (the glass spreader disc, Guild, 1931, as modified by Stiles, 1955, is only the most obvious) and imperfections in the experience and skill of our observers. There is no answer to such questions, but one pondering them may also recollect that even larger failures of radiance-invariance were found by Stiles whose skills as an observer, apparatus design, construction and meticulous care in attention to every detail of the experiment, have never been equaled.

The most surprising of these results is neither the failures of Grassmann's laws nor the ubiquitous frequency of their statistical significant incidence. What surprises (as testified by the results in Fig. 6, for example) is how small these deviations from radiance-invariance and additivity can be among trained normal trichromats! This is good evidence for the approximate (as distinct from the exact) validity of the laws of scalar multiplication and additivity and to the general utility of the pigment theory of matching though in details it is much too simple to deal with every complexity of real world colorimetry.

Acknowledgements—Assisted by a research grant EY 00197-31 from the NEI and by a Senior Scientific Investigator Award from Research to Prevent Blindness to M.A. We thank Fang Li and D. Wicker for many hours of devoted observation, A. Liepa for programming help, E. Arnold for discussions and Professor D. H. Krantz of Columbia University for numerous suggestions, criticisms and advice during several different stages of this work.

REFERENCES

Ahnelt, P. K., Kolb, H. & Pflug, R. (1987). Identification of a subtype of cone photoreceptor, likely to be blue sensitive, in the human retina. *Journal of Comparative Neurology*, *255*, 18–34.

Alpern, M. (1979). Lack of uniformity in colour matching. *Journal of Physiology, London 288*, 85–105.

Alpern, M. (1986). The Stiles–Crawford effect of the second kind (SCII): A review. *Perception*, *15*, 785–799.

Alpern M. (1989). The Charles Prentice Award lecture 1988: The directionality of color matches and its relation to secondary protanomalous trichromacy. *Optometry and Vision Science*, *66*, 339–354.

Alpern, M. & Tamaki, R. (1983). The saturation of monochromatic lights obliquely incident on the retina. *Journal of Physiology, London*, *338*, 669–691.

Alpern, M., Bastian, B., Pugh, E. N. & Gras, W. (1976). Altered ocular pigments photostable and labile. *Modern Problems in Ophthalmology*, *17*, 273–291.

Alpern, M., Kitahara, K. & Krantz, D. H. (1983a). Classical tritanopia. *Journal of Physiology, London*, *335*, 655–681.

Alpern, M., Kitahara, K. & Tamaki, R. (1983b). The dependence of the colour and brightness of a monochromatic light upon its angle of incidence on the retina. *Journal of Physiology, London*, *338*, 651–668.

Alpern, M., Kitahara, H. & Fielder, G. H. (1987). The change in color matches with retinal angle of incidence of the colorimeter beams. *Vision Research*, *27*, 1763–1778.

Blottiau, F. (1947). Les défauts d'additivité de la colorimétrie trichromatique. *Revue d'Optique théorique et instrumentale*, *26*, 193–201.

Brindley, G. S. (1953). The effect on colour vision of adaptation to very bright lights. *Journal of Physiology, London*, *122*, 332–350.

Brindley, G. S. (1960). *The physiology of the retina and visual pathways* (p. 211). London: Edward Arnold.

Cornsweet, T. N. (1970). *Visual perception* (p. 218). New York: Academic Press.

Crawford, B. H. (1965). Colour matching and adaptation. *Vision Research*, *5*, 71–78.

Enoch, J. M. & Stiles, W. S. (1961). The colour change of monochromatic light with retinal angle of incidence. *Optica Acta*, *8*, 329–358.

Guild, J. (1931). V. The colorimetric properties of the spectrum. *Proceedings of the Royal Society of London, B*, *230*, 149–186.

Krantz, D. H. (1975). Color measurement and color theory. I. Representation theorem for Grassmann structures. *Journal of Mathematical Psychology*, *12*, 283–303.

McCree, K. J. (1960). Colour confusion produced by voluntary fixation. *Optica Acta*, *7*, 281–290.

Pugh, E. N. Jr (1988). Vision: Physics and retinal physiology. In Atkinson, R. C., Herrnstein, R. J., Lindzey, G. & Luce, R. D. (Eds.), *Stevens' handbook of experimental psychology* (Vol. 1, Chapt. 2, pp. 116–118). New York: Wiley.

Rushton, W. A. H. (1972). Review lecture. Pigments and signals in colour vision. *Journal of Physiology, London*, *220*, 1P–33P.

Stiles, W. S. (1937). The luminous efficiency of monochromatic rays entering the eye pupil at different points and a new colour effect. *Proceedings of the Royal Society of London, B*, *123*, 90–118.

Stiles, W. S. (1955). The basic data of colour matching. *The physical society yearbook of 1955* (pp. 44–65). London: Physical Society.

Trezona, P. W. (1953). Additivity of colour equations. *Proceedings of the Physical Society*, *66B*, 547–556.

Trezona, P. W. (1954). Additivity of colour equations II. *Proceedings of the Physical Society*, *67B*, 513–522.

Trezona, P. W. (1973). The tetrachromatic colour match as a colorimetric technique. *Vision Research*, *13*, 9–25.

Trezona, P. W. (1974). Additivity in the tetrachromatic colour matching system. *Vision Research*, *14*, 1291–1303.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd edn, pp. 518–539). New York: McGraw-Hill.

Wyszecki, G. & Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae* (2nd edn, pp. 278–285, 379–392). New York: Wiley.

Zaidi, Q. (1986). Adaptation and color matching. *Vision Research*, *26*, 1925–1938.

APPENDIX

Tables A1 and A2 of this Appendix provide the statistical summaries of the Analysis of Variance of the metameric and isomeric matches respectively.

Table A1. Two-way (sessions and radiance) ANOVA mixed model metameric matches to $\lambda = 577.3$ nm (with nested random effects)

		Without rod-saturating background						With rod-saturating background					
Source	d.f.	MS(u)	F(u)	MS(v)	F(v)	MS(w)	F(w)	MS(u)	F(u)	MS(v)	F(v)	MS(w)	F(w)
HZ				[645.2 nm (15,500 cm^{-1}), 526.3 nm (19,000 cm^{-1}), 444.4 nm (22,500 cm^{-1})]									
Sessions	14	1.34E-2		1.60E-2		2.53E-4		1.22E-2		1.30E-2		3.67E-4	
Radiance	6	4.57E-4	6.83[c]	5.80E-4	8.75[c]	8.46E-6	10.03[c]	1.31E-3	9.52[c]	1.52E-3	11.56[c]	4.76E-6	3.04[b]
Radiance × sessions	84	6.68E-5		6.62E-5		8.44E-7		1.37E-4		1.32E-4		1.56E-6	
Radiance × runs (w sessions)													
FL				[645.2 nm (15,500 cm^{-1}), 526.3 nm (19,000 cm^{-1}), 444.4 nm (22,500 cm^{-1})]									
Sessions	9	1.12E-1	14.9	1.12E-1	15.61	2.03E-3	16.98	1.43E-1	13.9	9.59E-2	3.38	3.58E-2	1.6
Runs (within sessions)	10	7.51E-3		7.16E-3		1.20E-4		1.03E-2		2.83E-2		2.23E-2	
Radiance	6	1.58E-4	2.24[a]	3.63E-4	5.26[c]	1.18E-4	4.56[c]	1.93E-2	21.86[c]	1.4E-2	20.6[c]	2.96E-3	1.25[a]
Radiance × sessions	54	7.05E-5	0.64	6.9E-5	0.61	2.59E-5	0.92	8.84E-4	0.66	6.8E-4	0.74	2.37E-3	
Radiance × runs (w sessions)	60	1.10E-4		1.14E-4		2.80E-5		1.34E-3		9.14E-4		2.39E-3	
DW				[645.2 nm (15,500 cm^{-1}), 526.3 nm (19,000 cm^{-1}), 444.4 nm (22,500 cm^{-1})]									
Sessions	11	1.32E-2	4.05	1.32E-2	4.05	None used		1.65E-2	3.61	1.65E-2	3.61	None used	
Runs (within sessions)	12	3.25E-3		3.25E-3		None used		4.59E-3		4.59E-3		None used	
Radiance	6	1.53E-3	6.48[c]	1.53E-3	6.48[c]	None used		5.47E-4	1.19[a]	5.47E-4	1.19[a]	None used	
Radiance × sessions	66	2.37E-4	0.98	2.37E-4	0.98	None used		4.58E-4	0.85	4.58E-4	0.85	None used	
Radiance × runs (w sessions)	72	2.42E-4		2.42E-4		None used		5.40E-4		5.40E-4		None used	
HZ				[645.2 nm (15,500 cm^{-1}), 550.3 nm (18,172 cm^{-1}), 444.4 nm (22,500 cm^{-1})]									
Sessions	5	1.05E-2	13.63	8.60E-3	8.88	1.46E-3	2.04	9.60E-3	11.63	7.98E-3	7.97	1.50E-3	2.06
Runs (within sessions)	18	7.70E-4		9.69E-4		7.16E-4		8.25E-4		1.00E-3		7.28E-4	
Radiance	6	4.73E-4	59.33[c]	5.24E-4	61.34[c]	1.35E-6	30.08[c]	1.17E-3	40.31[c]	1.30E-3	40.37[c]	3.59E-6	26.93[c]
Radiance × sessions	30	7.97E-6	0.64	8.54E-6	0.65	4.50E-8	1.21	2.89E-5	1.92	3.21E-5	1.90	1.33E-7	1.46
Radiance × runs (w sessions)	108	1.25E-5		1.31E-5		3.73E-8		1.51E-5		1.69E-5		9.15E-8	
FL				[645.2 nm (15,500 cm^{-1}), 550.3 nm (18,172 cm^{-1}), 444.4 nm (22,500 cm^{-1})]									
Sessions	5	6.17E-2	20.7	6.07E-2	17.23	8.02E-4	2.32	5.98E-2	15.61	5.99E-2	12.58	7.79E-4	2.08
Runs (within sessions)	18	2.98E-3		3.52E-3		3.46E-4		3.83E-3		4.76E-3		3.74E-4	
Radiance	6	2.48E-4	11.12[c]	2.64E-4	10.92[c]	5.38E-6	1.37[a]	1.82E-3	8.06[c]	1.94E-3	8.06[c]	1.94E-6	7.47[c]
Radiance × sessions	30	2.23E-5	0.72	2.42E-5	0.73	3.92E-6	0.96	2.26E-4	1.04	2.41E-4	1.05	2.60E-7	1.35
Radiance × runs (w sessions)	108	3.10E-5		3.32E-5		4.10E-6		2.17E-4		2.29E-4		1.92E-7	

[a] = F not statistically significant; [b] = F significant at least the 0.05 level, [c] = F significant at least the 0.001 level.

Table A2. Two-way analysis of variance of scalar multiplication (mixed model) isomeric matches to $\lambda = 577.3$ nm

		Without rod-saturating background						With violet rod-saturating background					
Source	d.f.	MS(u)	F(u)	MS(v)	F(v)	MS(w)	F(w)	MS(u)	F(u)	MS(v)	F(v)	MS(w)	F(w)
HZ						(Open circles in Fig. 2)							
Sessions	5	9.81E-5	21.27	9.34E-5	20.54	5.41E-8	33.25	1.20E-4	8.10	1.17E-4	7.93	4.10E-8	6.93
Runs (without sessions)	18	4.61E-6		4.55E-6		1.63E-9		1.48E-5		1.48E-5		5.91E-9	
Radiance	6	2.15E-5	5.93[c]	2.07E-5	5.81[c]	3.02E-9	4.27[b]	3.88E-6	0.86[a]	3.79E-6	0.85[a]	1.49E-9	3.57[b]
Radiance × sessions	30	3.63E-6	2.29	3.57E-6	2.34	7.06E-10	0.95	4.51E-6	1.64	4.46E-6	1.63	4.17E-10	0.57
Radiance × runs (w sessions)	108	1.59E-6		1.52E-6		7.47E-10		2.75E-6		2.73E-6		7.28E-10	
FL						(Open triangles in Fig. 2)							
Sessions	9	3.13E-4	12.53	3.07E-4	12.71	2.65E-8	13.25	8.40E-4	2.71	8.28E-4	2.72	7.02E-8	2.76
Runs (within sessions)	10	2.50E-5		2.41E-5		2.00E-9		3.10E-4		3.04E-4		2.54E-8	
Radiance	6	7.41E-5	6.26[c]	7.27E-5	6.26[a]	6.41E-9	6.41[c]	1.19E-3	13.75[c]	1.16E-3	13.71[c]	9.84E-8	13.98[c]
Radiance × sessions	54	1.18E-5	0.97	1.16E-5	0.97	1.00E-9	0.99	8.65E-5	1.41	8.49E-5	1.40	7.04E-9	1.39
Radiance × runs (w sessions)	60	1.22E-5		1.20E-5		1.01E-9		6.14E-5		6.05E-5		5.07E-9	
DW						(Not illustrated)							
Sessions	11	1.47E-3	14.78	1.47E-3	14.78	None used		4.65E-3	16.22	4.65E-3	16.22	None used	
Runs (within sessions)	12	9.95E-5		9.95E-5		None used		2.86E-4		2.86E-4		None used	
Radiance	6	1.89E-5	1.40[a]	1.89E-5	1.40[a]	None used		3.76E-5	0.79[a]	3.76E-5	0.79[a]	None used	
Radiance × sessions	66	1.36E-5	0.94	1.36E-5	0.94	None used		4.76E-5	0.87	4.76E-5	0.87	None used	
Radiance × runs (w sessions)	72	1.45E-5		1.45E-5		None used		5.45E-5		5.45E-5		None used	

[a] = F not statistically significant; [b] = F significant at least the 0.01 level; [c] = F significant at least the 0.0001 level.