Vision Res. Vol. 14, pp. 1291 to 1303. Pergamon Press 1974. Printed in Great Britain

ADDITIVITY IN THE TETRACHROMATIC COLOUR MATCHING SYSTEM

P. W. TREZONA
National Physical Laboratory, Teddington, England[1]

(*Received* 20 *February* 1974; *in revised form* 30 *April* 1974)

Abstract—The trichromatic system of measurement, although suitable for fields subtending around 2°, does not give a linear metric with larger fields because of the presence of rod receptors. Their presence is allowed for in the tetrachromatic colour match and so one might expect the tetrachromatic system to be more linear. It is the purpose of this paper to test this by comparing tetrachromatic and trichromatic additivity for a variety of test stimuli. A special case of the additivity principle, luminance metamerism, presents a much more severe test than the general case: while the latter is only touched on, the former is concentrated on. Most experiments involve measurement but some employ subjective judgement. No departure from tetrachromatic additivity has yet been detected, although trichromatic nonadditivity is pronounced. Trichromatic additivity failures so far examined are always in a direction consistent with the hypothesis that "rod colour" is blue. An indication is given of the way in which the additive property of the tetrachromatic colour match can be utilized to develop general systems of colorimetry and photometry, applicable at all luminance levels.

INTRODUCTION

The basic criterion by which a set of colour matching functions[2] should be judged is the extent to which they are subject to the laws of additivity. Without the validity of these laws, colour matching functions are of no wider application than under the experimental conditions of their measurement, so that colour matching measurements cannot be assumed to hold at other luminance levels, they cannot be combined and they cannot validly be linearly transformed to another set of matching stimuli, either real or imaginary. The possibility of imaginary stimuli is important in allowing choice either for easy manipulation or for physiological significance. Successful industrial applications of colour matching functions for, say, computing the chromaticity of a colour filter when its spectral transmission characteristics and the spectral power distribution of the source are known, is limited by the extent to which additivity is valid: the same can be said of instruments with responses which are designed as linear transformations of the colour matching functions.

Additivity for centrally viewed 2° fields has been tested by Blottiau (1947) and Trezona (1953, 1954) and although some deviations were found that were probably experimentally significant, the procedure can be considered valid for all practical purposes. But a field of 2° subtense is unrepresentative of many real viewing situations and it can in no way be regarded as typical of the retina as a whole. The only rod free area of the retina occupies most of the central 2° location to which the 1931 C.I.E. Colorimetric System refers, and this is also the region of the Maxwell Spot phenomenon. With a larger field judgement will be made on some combination of fovea and periphery: relative to its size the former will have the greater importance, but in the 10° field the area containing rods is more than 96 per cent of the whole. A consequence of rod activity in large fields is non-additivity, first shown by Stiles (1955)—Fig. 1. This shows how actual matches at various levels depart from those predicted from colour matching functions measured at high levels.

In spite of this inherent defect in large field trichromatic colorimetry, but because of industrial need, the C.I.E. in 1964 adopted the large field (10°) colour matching functions as standards. These became the basis of the C.I.E. supplementary standard observer, adopted in addition to the 1931 standard observer which had been based on a 2° bipartite field. The C.I.E. recommended the use of the large field standards whenever the field size exceeds 4° subtense. These large field standards were based mainly on the work of Stiles (1955, 1958) and to a lesser extent on that of Speranskaya (1958, 1959), both using a 10° field. However a note of caution has since crept in and the C.I.E. Publication No. 15 (E-1.3.1) 1971 states "The large-field colour-matching data as defined by the C.I.E. 1964 supplementary standard colorimetric observer are intended to apply to matches where the luminance and the relative spectral power distributions of the matched stimuli are such that no participation of the rod receptors of the visual mechanism is to be expected. This condition of observation is important as 'rod intrusion' may upset the predictions of the standard observer". This recommendation poses a problem

---

[1] Now at the City University, London.

[2] An alternative name is "spectral tristimulus values".

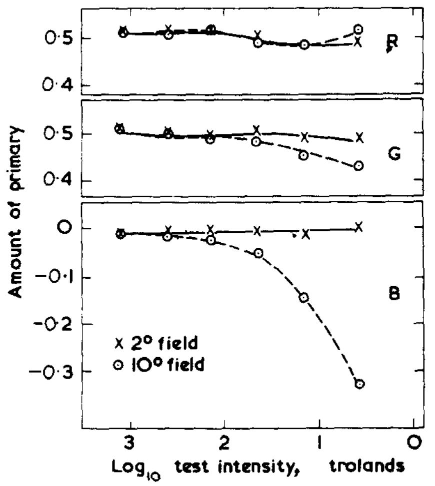

Fig. 1. The effect of changing the field luminance on the matching stimuli proportions for a match on wavelength 581 nm. No match change is indicated by a horizontal line (after Stiles, 1955).

as it is very difficult to know under what conditions rods are not participating. Aguilar and Stiles (1954) showed that under certain circumstances rod activity can persist up to high retinal illuminations of around 2000 scotopic td. Wyszecki (1972) stated "In large-field viewing situations the rod mechanism may not completely be eliminated from functioning even at very high luminance levels. When strong metamerism is present, particularly for red colors, the intrusion of the rods may be significant and make match-predictions by a trichromatic system (such as the C.I.E. large field system) incorrectly".

To summarize the situation, in both conventional and large field colorimetry trichromatic matches can be satisfactorily performed. However, whereas the former lead to a linear metric, the latter do not. It is fairly certain that the reason for this is match distortion caused by the presence of rods in all fields larger than 2°. For a linear metric with large fields a different kind of colorimetry is required: the presence of rods introduces an extra degree of freedom, necessitating a tetrachromatic rather than trichromatic system of measurement.

The tetrachromatic colour match was first attempted by Bongard *et al.* (1958) using a trial and error technique, and indeed a number of points on the colour matching functions were measured, but the work was not pursued. In 1970 two systematic methods were proposed, by Palmer (1970) and Trezona (1970). Palmer's method was an indirect one, basing the tetrachromatic match setting on the point of intersection of linear graphs representing certain previous matches, and several tetrachromatic colour matches by this method were later carried out (Palmer, 1972).

Trezona's proposal was a direct method of reaching the tetrachromatic match by using the two new concepts of making a match to satisfy the observer at two luminance levels and an approach to the final match setting by iterative convergence. Later the technique of making this match was developed and certain properties were studied (Trezona 1972, 1973a). It was shown to be a unique match, independent of the starting conditions: Fig. 2 shows how, for each matching stimulus, even for widely differing starting conditions, there is a convergence towards the final tetrachromatic match setting. This uniqueness was shown to be general for all spectral and non-spectral colours tried. Even under the most adverse starting conditions, the tetrachromatic match can be reached in about 15 min: this is fast enough for it to be used for a series of measurements within an observing session of reasonable duration. The coloured Maxwell Spot, very disturbing for some people in large field trichromatic colour matching, has never been observed in the tetrachromatic match (Trezona 1973a).

Since the tetrachromatic match equates rod activity as well as that of the three chromatic cone processes, one would predict that additivity deviations caused by rod discrepancy between the two fields in the trichromatic system, should be abolished. The present paper is designed to test just this. In order to put additivity tests in perspective, tetrachromatic and trichromatic tests are both made on a given test stimulus. Since match discrepancy cannot be isolated from match imprecision, the question of precision is considered in all tests, trichromatic and tetrachromatic.

Additivity tests involving measurement are of two types. Equations given below are appropriate to the trichromatic system but with an extra term they would also apply to the tetrachromatic.

The general form of the additivity principle can be stated as:

If

$$dD = r_1\mathrm{R} + g_1\mathrm{G} + b_1\mathrm{B} \qquad (1)$$

and

$$eE = r_2\mathrm{R} + g_2\mathrm{G} + b_2\mathrm{B} \qquad (2)$$

then

$$dD + eE = (r_1 + r_2)\mathrm{R} + (g_1 + g_2)\mathrm{G} + (b_1 + b_2)\mathrm{B}. \qquad (3)$$

These equations refer to R, G, B colour matches on colours D and E separately and in combination and in general dD and eE may be different in both chromaticity and luminance.

A special case is when the component colours have the same chromaticity. Adding n such equations as (4) leads to (5), predicting that the colour match should be invariant with respect to luminance level.

$$dD = r\mathrm{R} + g\mathrm{G} + b\mathrm{B} \qquad (4)$$

$$ndD = nr\mathrm{R} + ng\mathrm{G} + nb\mathrm{B}. \qquad (5)$$

Testing of the visual validity of both forms (general and special) is necessary. However, most tests in this paper were made on the latter as this is a more specific and hence more suitable test for any non-additivity caused by rod participation (see Discussion).

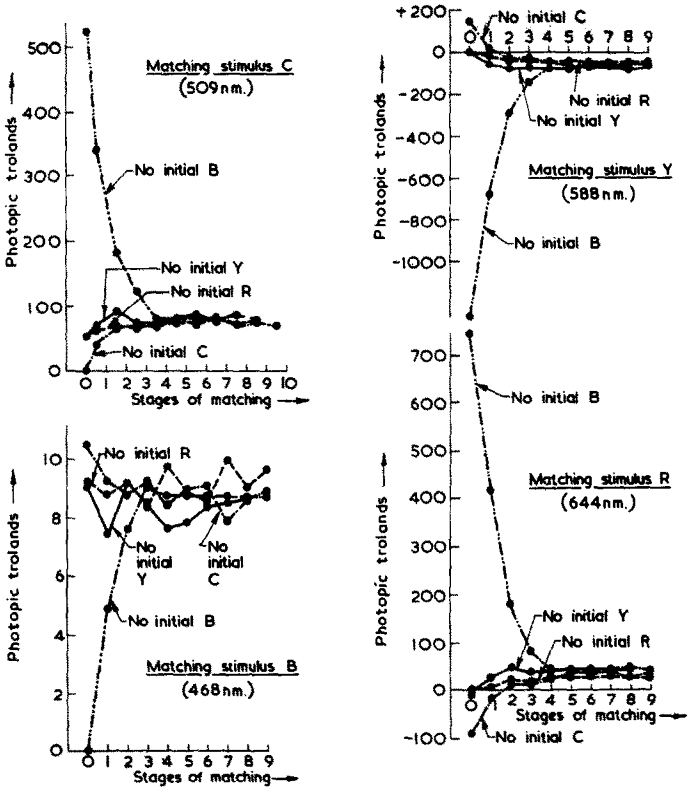

Fig. 2. Approach to the tetrachromatic colour match from the four extreme starting conditions. Amount of each matching stimulus required for a match on test stimulus 485 nm (retinal illumination 55 photopic td) is plotted against the number of steps taken towards the final match. Stage O is the initial trichromatic match. Low level is 1/720th of experimental level. Observer PGM.

Figure 1 shows Stiles' (1955) results of how a trichromatic colour match changes with luminance level for a 2° and a 10° field. Although no appreciable change occurs for the 2° field, the change is large in the 10° case: this effect is referred to as *luminance metamerism*. Results of many additivity and related tests on large fields were also disappointing (Clarke, 1960, 1963; Wright and Wyszecki, 1960; Stiles and Wyszecki, 1962; Stiles, 1963; Nimeroff, 1964a, 1964b; Crawford, 1965; Lozano and Palmer, 1967, 1968). The preponderance of rods in fields subtending more than 2° changes the magnitude of the problem of additivity deviations from the practically negligible one of the 2° field into the highly significant one of the 10° field, severely limiting the use of a large field trichromatic system.

METHOD

General remarks

Measurements were made on the NPL Trichromator (Stiles, 1955) modified to become a tetrachromator (Clarke, 1973; and Clarke and Trezona, to be published). The technique of tetrachromatic colour matching used was as described by Trezona (1972, 1973a) and the same matching stimuli were used, i.e. R (644 nm), Y (588 nm), C (509 nm) and B (468 nm). Briefly, the technique is to make matches at two luminance levels, one being the level of the experiment and the other a "low level", below cone threshold. While the former is a trichromatic match using R, Y and B, the latter is an achromatic brightness match, one stimulus (C) sufficing. Each match upsets the previous one, but after a few iterations a match is reached which satisfies the observer at both levels: this is the tetrachromatic colour match.

There were three observers, PGM, ERF and RDL, all male and all having normal colour vision. Unlike the other two, observer PGM never sees the Maxwell Spot. All could be regarded as trained observers although RDL had had no previous experience of tetrachromatic colour matching. Positioning of the eye's pupil was by a telescope (Stiles, 1955) but observer ERF appears to have a decentered pupil and so for him only the vertical plane adjustment was made in this way: adjustments within this plane used small field chromatic aberration. Observing sessions were preceded by 10 min dark adaptation. Each session lasted 1·5–2 hr; sometimes the observer required a short rest and then dark adaptation again took place before the second part of the session.

In the main, tests were concerned with measurement, although one subjective judgement experiment was done.

1. *Measurement tests*

In a given observing session matches, sometimes trichromatic and sometimes tetrachromatic, were made under four (sometimes three) different conditions. In this way each complete set of conditions investigated were covered on the same occasion.

In order to avoid the extra burden of reaching the tetrachromatic match from scratch, a preliminary experiment had been done on a previous occasion to find the approximate match setting: in each case the more favourable starting conditions had been used, i.e. those avoiding a change of sign of any matching stimulus (Trezona, 1972, 1973a). Then on the day of the experiment six matches were made, each consisting of a low and experimental level component, performed in that order. The first match was discarded, leaving five readings for each set of conditions: by doing this any day-to-day visual change that might have occurred was catered for. Although, strictly, successive tetrachromatic matches are not truly independent, in view of the large *possible* successive changes, they may be regarded as such (Trezona, 1973a).

Measurement experiments fell into two categories. The major investigation tested for any match change with luminance level, i.e. luminance metamerism, the special case (see above), while the minor investigation tested an example of the general case.

(a) *Special case of additivity test.* Four luminance levels were presented at random, these differing by × 6 luminance steps. Calling the highest level I, levels $I/6$ and $I/6^2$ were obtained by suitably reducing the luminance of the test stimulus. Level $I/6^3$ could have been similarly produced, but for convenience (especially to avoid frequent low level and hence more noisy photomultiplier measurements) level $I/6$ was used in conjunction with a rotating sector disc of total angle 10°, run at above the critical fusion frequency of flicker. In subsidiary experiments using each of two different test stimuli, PGM checked that at level $I/6^3$ no statistically significant difference between the two methods can be found either in the trichromatic or the tetrachromatic case.

Test stimuli were in turn 530, 610 and 546 nm, the investigation being performed by PGM. Repeatability was tested when PGM again used 530 nm, and observer change was tested when RDL did the experiment on 546 nm. This made five investigations in all.

For each test stimulus the tetrachromatic investigation was made and either one or two trichromatic investigations. For trichromatic tests, R (644 nm) and B (468 nm) were always used and either Y (588 nm) or C (509 nm). For 530 nm, Y was used as C did not show large additivity deviations and the reverse may be said of 610 nm. To anticipate the criticism that the tetrachromatic system is superior to the trichromatic only because it, in effect, allows a choice designed to give the best results, between Y and C, a test stimulus wavelength had been looked for which gave non-additivity in *both* trichromatic systems. PGM had found 546 nm to be such a wavelength: RDL used the same wavelength without a preliminary test in his case.[3]

An experiment associated with the above, and planned to confirm results using a different method, was undertaken by the third observer ERF. With test stimulus 546 nm and in turn trichromatic (R, Y, B), trichromatic (R, C, B) and the tetrachromatic case, a match was made at level $I/6$ and then viewed at level $I/6^3$ using the sector. At this low level he was asked to judge whether the match still held and if not to attempt to restore the match using B only. Judgement was then made as to whether the match was completely restored and all quantities of B were recorded.

Next the tetrachromatic match was viewed at all levels between the experimental level and the region of the threshold of vision in × 2 steps (20 levels in all) to see if a mismatch could be detected.

(b) *General case of additivity test.* A low level additivity test was made on test stimuli 546 nm and 478 nm by observer PGM. They were matched in combination and then separately, so that the predicted sum could be computed and compared with the measured mixture. This was done for the tetrachromatic and trichromatic (R, Y, B) system.

2. *Subjective judgement test*

Observer ERF extended the investigation to all five spectral regions using a completely different technique. From each of the five spectral regions that occur with four matching stimuli, a test stimulus was chosen. For each one a tetrachromatic match was made and then all four possible trichromatic matches for a certain test stimulus luminance level. Then each was viewed at 0·05 × this level and the mismatch assessed on a 1–4 subjective judgement scale, taking note of which field appeared brighter.

The purposes of this experiment were: (a) to extend the investigation to all spectral regions and all trichromatic matching stimuli combinations; (b) to test additivity using an entirely different technique; (c) to relate any changes directly to their visual significance, since it cannot necessarily be assumed that quite a large *measured* change is *perceived* as different by the observer.

RESULTS

1. *Measurement tests*

(*a*) *Special case of additivity test.* Figure 3 shows the effect of decreasing the luminance level on both tetrachromatic and trichromatic colour matches with test stimulus 530 nm. There are four graphs, one for each of the red (R or DR), yellow (Y or DY), cyan (C or DC) and blue (B or DB) matching stimuli. The luminance level decreases left to right on a log scale. A horizontal line indicates no match change with level: if a change occurs the line will depart from the horizontal. Since the particular trichromatic system chosen is (R, Y, B) there is no trichromatic graph for C. Each point represents the mean of five readings: any changes must be judged in conjunction with match uncertainty. To be consistent with reaching the match by the bracketing technique on a density wedge and with the eye's discrimination properties, the mean M, the standard error of the mean E and the limits $M \pm 2E$ were computed logarithmically. However, since it is powers and not their logarithms that are being tested for additivity, means and limits were then expressed linearly in photopic trolands and shown as such on the graphs. Another set of five readings would have a 95 per cent probability of yielding a mean which lies within these limits, but only if the eye could be regarded as a completely unchanging system merely suffering from imprecision. Since this is not so, results must be interpreted with care. In Fig. 3 changes in the red, yellow

---

[3] Observer RDL was only available for a few hours during a short visit from abroad.

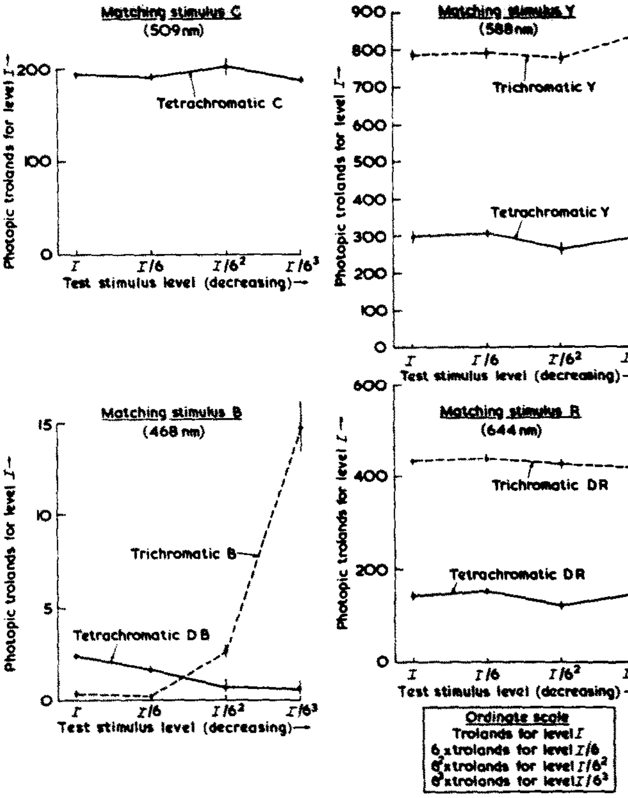

Fig. 3. Trichromatic and tetrachromatic colour matches at four luminance levels for test stimulus 530 nm, of retinal illumination 400 photopic td at experimental level *I*. Retinal illumination of the field at level *I* is about 800 photopic td in the trichromatic case and 500 in the tetrachromatic. Low level = 1/2160th of experimental level *I*. Lines show ±2 × standard error of the mean. Observer PGM, first run.

and cyan matching stimuli are probably not significant: but the tetrachromatic DB shows a small decrease with decreasing level while the trichromatic B shows a large increase. The former occurs in the region of the blue threshold and is instrumental in cause: Trezona (1954) found similar blue changes with small fields (see Discussion). However increases in B cannot be explained in operational terms, and the trichromatic change must be regarded as an additivity breakdown of considerable magnitude.

In a repeat on a different occasion by the same observer (Fig. 4) the same pattern emerges. This time the trichromatic B breakdown is even more pronounced. The use of Figs. 3 and 4 together is useful as it confirms that small changes such as those of C at level $I/6^2$ are not significant, being in opposite directions on the two occasions.

Roughly the same remarks can be made concerning test stimulus 610 nm (Fig. 5) although some other small changes may also be present.

Results for PGM using test stimulus 546 nm, chosen by him to show non-additivity in both trichromatic systems, can be seen in Fig. 6. The tetrachromatic system is again additive for this test stimulus which shows non-additivity in both trichromatic systems. Tests on the same test stimulus by another observer RDL (Fig. 7) show tetrachromatic additivity and trichromatic (R, Y, B) non-additivity for the blue matching stimulus. Trichromatic (R, C, B) changes are probably not significant for this observer for whom the test stimulus 546 nm was not specifically selected. RDL's error lines for tetrachromatic colour matching are considerably longer than PGM's although his trichromatic ones are comparable. This is in keeping with the fact that RDL was skilled in trichromatic colour matching but had never before attempted tetrachromatic matching.

In Figs. 3–6 inclusive it can be seen that precision, as indicated by the shortness of error lines, is rather worse in the tetrachromatic than the trichromatic case, Fig. 7 being excluded for the reason given above. In making this comparison it must be remembered that for the same intrinsic error, the length of the error line might be expected to be proportional to the luminance level because of the Weber–Fechner Law which probably applies to these conditions. However, tetrachromatic precision is still good enough to consider the tetrachromatic match as a precise colorimetric technique.

In the associated experiment by the alternative tech-

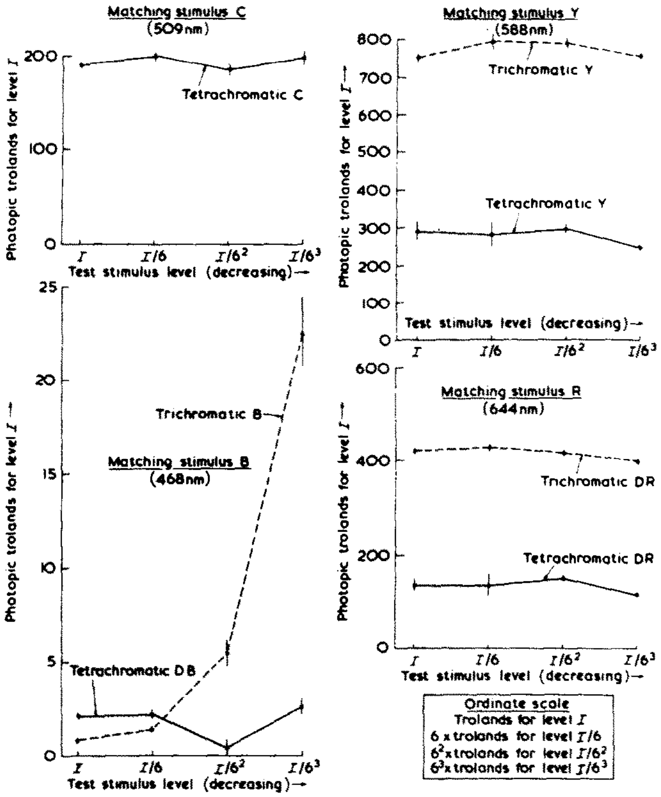

Fig. 4. As Fig. 3, second run.

nique, ERF's viewing of level $I/6^3$ trichromatic matches made on 546 nm at level $I/6$ for (R, Y, B) and (R, C, B) systems, showed that at this level they were quite unacceptable as matches. But the match could be completely restored in each case only by increasing the amount of the blue matching stimulus. This is in spite of the fact that the match failure was *described* in the former case as having a magenta test field and a dull yellow mixture field. In the (R, Y, B) system the increase in quantity of blue matching stimulus was × 13, compared with PGM's × 20 for the complete rematch. In the (R, C, B) system it was × 4 compared with PGM's × 2·5.

The tetrachromatic match test at 20 different levels showed that this match held in every case.

It should be noted that the colour itself changes with luminance level in the tetrachromatic as well as the trichromatic system: but in the former, unlike the latter, the match holds good.

(*b*) *General case of additivity test*

Figure 8 shows for each matching stimulus how the predicted sum compares with the measured mixture for both trichromatic and tetrachromatic cases: the two component matches in each case are also shown. Since it was the trichromatic (R, Y, B) system, there is no trichromatic C. Desaturating stimulus quantities are shown as negative.

Agreement between predicted sum and measured mean was good for all four matching stimuli in the tetrachromatic case. For the trichromatic, there was a pronounced discrepancy for B, the predicted value being larger than the measured: possibly there is a small R discrepancy.

The former discrepancy (but not the latter) was confirmed in a different way by ERF. For the trichromatic system only, he made the component matches and was presented with the predicted sum for appraisal. It was an unacceptable match but could be completely restored by decreasing B only.

2. *Subjective judgement test*

Table 1 shows results for ERF on a four point scale where assessment 4 indicates no mismatch. There is no luminance metamerism in the tetrachromatic case for any spectral region, but for the trichromatic system some matching stimulus combination shows it in every test stimulus region. The extent of luminance metamer-

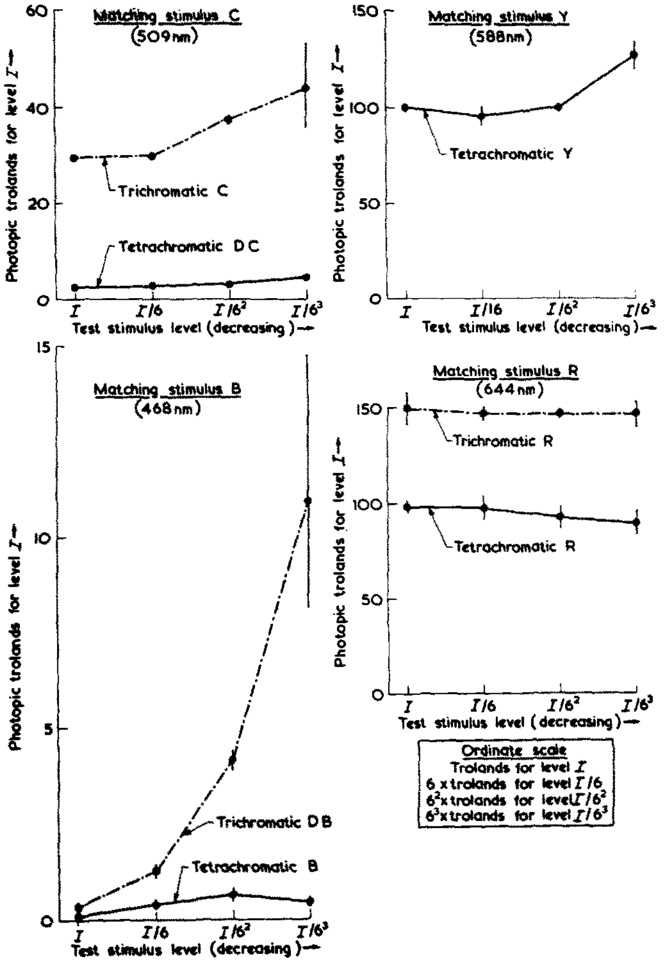

Fig. 5. Trichromatic and tetrachromatic colour matches at four luminance levels for test stimulus 610 nm, of retinal illumination 176 photopic td at experimental level *I*. Retinal illumination of the field at level *I* is about 200 photopic td in the trichromatic case and 180 in the tetrachromatic. Low level = 1/1440th of experimental level *I*. Lines show ±2 × standard error of the mean. Observer PGM.

ism in the different regions tends to be either uniformly indifferent or small for some combinations while large for others.

DISCUSSION

General and special cases

Although at first sight the general test of additivity, equations (1)–(3), would appear to be more important than any special case (see Introduction), it is in fact the luminance metamerism test that is a better one if rod intrusion is, indeed, the cause of additivity failures found by various workers. This is because the mode of action of rods is to participate to an increasing extent as the luminance level is reduced.

In the general case, quantities of a matching stimulus on occasions will subtract rather than add and so any scotopic input into the same channel will also subtract: hence the effect on measurements of any non-additivity caused by rod intrusion may be minimized. In fact for every stimulus pair, unless both come from the same one of the five spectral regions, subtraction rather than addition will occur for some matching stimuli. This indicates that the two combined colours should at least be close in wavelength, if not identical. The test stimuli in Fig. 8, 478 nm and 546 nm were chosen to be in the same trichromatic (R, Y, B) region and so all trichromatic quantities add; however they are not in the same tetrachromatic region and quantities of matching stimuli R, Y and B subtract, indicating that in this case it is a poor test for tetrachromatic additivity. In the luminance metamerism test, quantities always add.

Leaving this aside, there is an even more important difference. In the general case only two wavelengths are combined, but where the luminance level is changed by a factor of 6^3, this is equivalent to adding

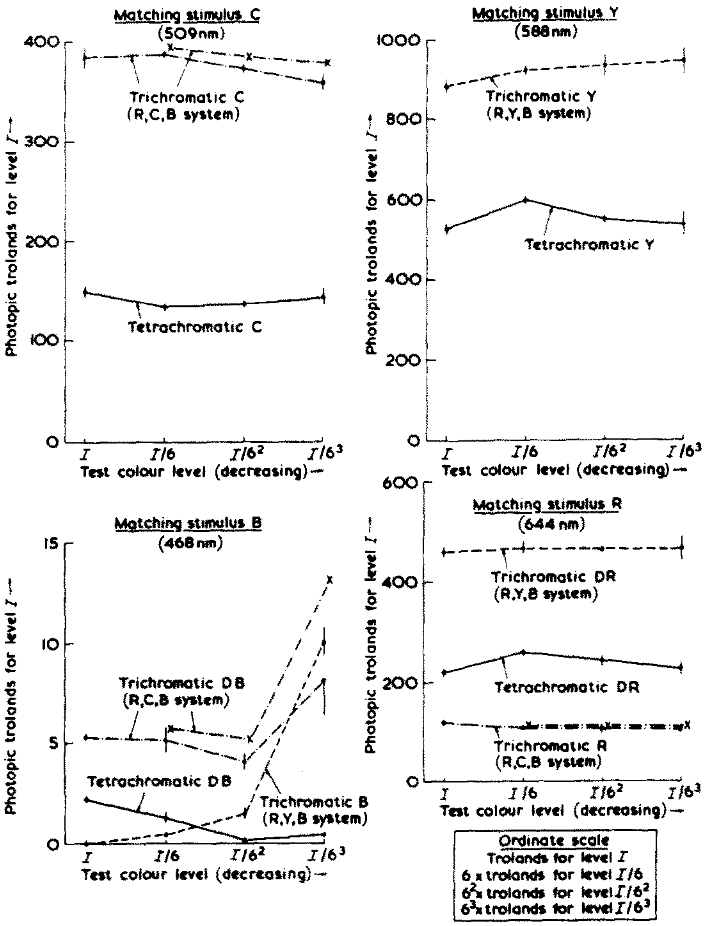

Fig. 6. Trichromatic (R,Y,B), trichromatic (R,C,B) and tetrachromatic colour matches at four luminance levels for test stimulus 546 nm, of retinal illumination 520 photopic td at experimental level *I*. Low level = 1/1200th of experimental level *I*. Lines show ±2 × standard of the mean. × Repeat at slightly lower level for trichromatic (R,C,B). Observer PGM.

216 separate (identical) colours: furthermore less uncertainty of measurement is involved since this luminance metamerism test involves only two colour matches compared with the three of the general case. Thus the luminance metamerism test is a more severe test of additivity than the general case. It is also more important because it is a fact that large variations in illumination occur in practical viewing of colour matched products, yet the consumer will expect the matches to hold.

This accounts for the concentration of effort on variation of luminance tests rather than on the general test of additivity. More work on this latter would probably have been done had the experiment involved a simpler technical procedure, but the modifications to the NPL instrument which would have been necessary to make it straightforward did not seem to be warranted in view of the above. For the test shown in Fig. 8 where wavelengths were chosen as discussed above, in order to make the effect appreciable it was also necessary to work in the lower mesopic range of luminance levels, where change of scotopic contribution with level is likely to be greatest. It may be necessary to develop a special technique to test for non-additivity in the general case, such as interleaving single matches for each component and their combination.

Considering the luminance metamerism tests, two methods were used, i.e. measurement and subjective judgement, the former giving quantitative exactitude and the latter being a better indication of what is actually seen and is also quick to perform. However the latter method, whereby a match is made at one level and inspection at another must be used with care since two kinds of error can be made. One kind of error is where a match is accepted because it lies within a just discriminable step but is displaced from its centre, i.e. it would not have been reached by the bracketing technique especially if the mean of several

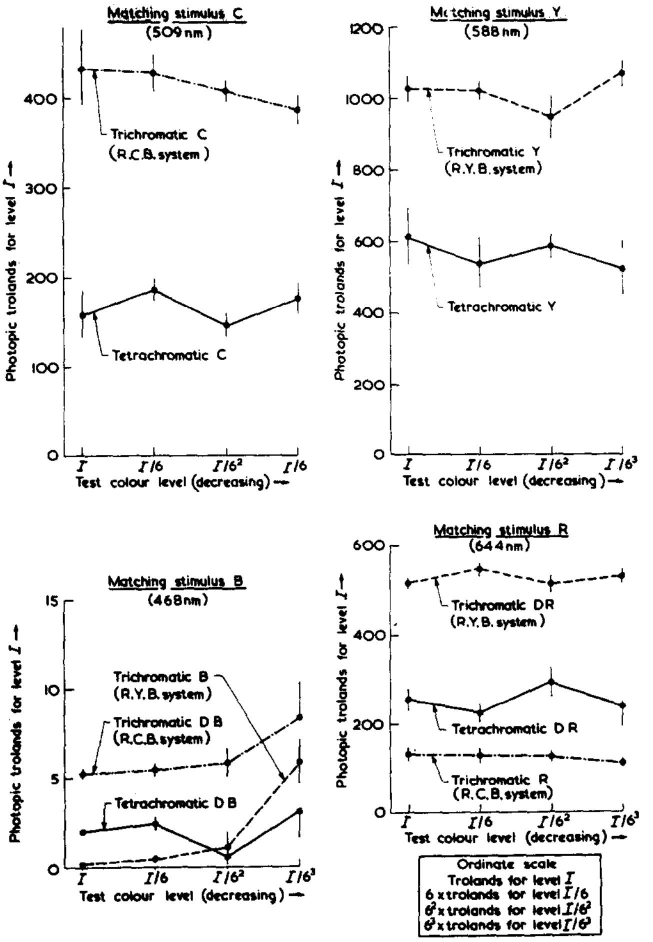

Fig. 7. As Fig. 6, observer RDL.

settings were used. The other kind of error, where a match is rejected even though additivity failure is not occurring, could happen if matching at some levels were more precise than at others and a match made at a less precise level were viewed at a more precise level: the "wrong" rejection would happen whenever the former match happened to differ appreciably from the mean of several settings. However, this "wrong" rejection at the worst is likely to apply to assessment 3 in Table 1 and certainly not to assessments 1 and 2.

MISMATCH FOR LUMINANCE × 0·05
FOR 10° FIELD

Test wavelength (nm)	RI (photopic trolands)	TRICHROMATIC No R	No Y	No C	No B	Tetra-chromatic
440	2·5	3 M	2 M	2 M	3 M	4
485	10	3 T	3 T	2 T	3 M	4
530	62	3 M	3 M	1 T	2 T	4
610	65	1 T	1 M	4	4	4
700	24	1 T	1 T	4	1 M	4

R 644 nm
Y 588 nm
C 509 nm
B 468 nm

T Test field brighter
M Mixture " "

1 Mismatch pronounced
2 " distinctly visible
3 " just visible
4 No mismatch (or uncertain)

Observer E.R.F.

The test conditions selected to compare any tetrachromatic and trichromatic non-additivity were specifically chosen where the latter was known to be bad. However, it is not possible to choose conditions on the basis of failures of the former, since it is not known where tetrachromatic non-additivity is likely, if indeed it occurs at all.

What restrictions are permissible in a system of colorimetry?

In a system of trichromatic colorimetry which allows linear transformations, *any set* of three matching stimuli which give sufficient precision should be allowed. Even if this is not fully accepted, one might expect a *single set* of three matching stimuli to be used for all test stimuli and not chosen as, say, (R, C, B) for

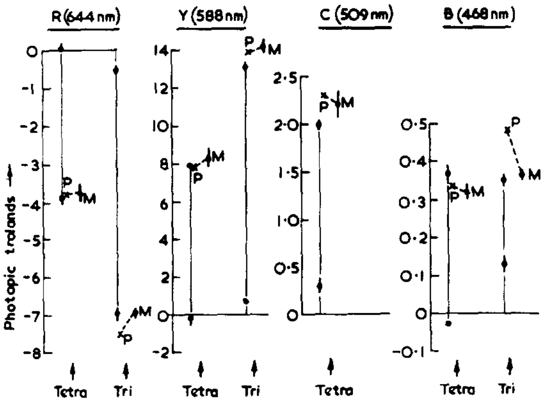

Fig. 8. Additivity test on 478 nm and 546 nm for trichromatic and tetrachromatic colour matches. Retinal illumination of 546 nm is 7 photopic td and of 478 nm is 0·6. Components of the algebraic sum are shown separately (•) and combined (×). The Measured Mixture (*M*) is joined to the Predicted Sum (*P*) by dotted line. Lines show ±2 × standard error of the mean. Observer PGM.

blues and greens and (R, Y, B) for oranges and reds in order to avoid luminance metamerism: Figs. 3, 4 and 5 show that wavelength 530 nm gives luminance metamerism for one trichromatic system (R, Y, B) and 610 nm for another (R, C, B). But even if this argument is not subscribed to it and it is claimed that in the trichromatic system a *choice between Y and C* should be allowed according to the spectral region, then a comparison between the tetrachromatic and the *better* trichromatic system must be made: Fig. 6 shows that the tetrachromatic system is clearly superior to either trichromatic system for some stimuli.

Apparent additivity deviations, having an operational cause

It has been mentioned that a *decreased* quantity of B compared with that predicted can occur near the threshold of colour for blue due to instrumental reasons and was also found for the additive or nearly additive 2° field work (Trezona, 1953, 1954). Levels $I/6^2$ and $I/6^3$ (Figs. 3 and 4) both lie in the doubtful region around the threshold of colour for B on its own, and it is around here that the decreases occur. The normal technique of colour matching, whereby each matching stimulus is set centrally between the extremes of bracketing on a density wedge, successfully locates a quantity X as the mean, to the extent to which the threshold increments $\pm\ \Delta \log X$ are equal. Although this is never grossly wrong for most colours and at most levels, near the threshold $-\Delta \log X$ becomes very large compared with $+\Delta \log X$ and the centre of bracketing will be located downwards from the quantity X (Trezona, 1954). Any *increase* at low levels can never be explained by such an argument and indeed its cause may well be of such a magnitude as to override what would otherwise tend to be a decrease.

Stability and precision

Returning to Figs. 3 and 4, it is interesting that all trichromatic B values on the second occasion are nearly double those on the first. Such day-to-day instability in B readings has been noticed on other occasions (e.g. Trezona, 1973a).

These and other figures also indicate that tetrachromatic precision is worse than trichromatic, although not unduly so. The cause is the lower precision of the low level component of the tetrachromatic match. However, all four matching stimuli are less precise and not just C which is used at the low level. Suppose on one occasion C is displaced downwards from the mean by an amount which would be large by trichromatic standards: then at the experimental level the decrease in C must be counteracted by increases of R, Y, B to just compensate for this. On another occasion when C is displaced upwards, R, Y and B must be decreased to compensate: hence all matching stimuli are imprecise to the same extent.

Non-additivity of the blue matching stimulus related to the scotopic discrepancy between the fields

Calculations based on the C.I.E. $V'(\lambda)$ curve indicate that for test stimulus 546 nm at level *I*, Fig. 6, and the trichromatic (R, Y, B) system the test field has a scotopic luminance equal to 3·5 × that of the mixture field. If one makes the assumption that "rod colour" is blue (Trezona 1970) one would predict an increase of B to compensate in the mixture field as luminance level decreases and hence relative rod activity increases: this is in fact what happens. Similar calculations for the trichromatic (R, C, B) system show that this time the mixture field has the greater scotopic luminance, being 2·5 × that of the test field. Now one would predict a compensating increase of DB in the test field with decreasing luminance level: again this occurs. The smaller ratio in the second case might perhaps lead one to expect a smaller degree of the effect, which does in fact happen. At level I for the tetrachromatic case, the scotopic luminances of the fields for the various matches agree to a standard deviation of order 5 per cent: however the C.I.E. $V'(\lambda)$ curve based on a 20° field was used, rather than the more applicable scotopic function of the actual observer under the 10° field conditions. Hence one is lead to expect no change of DB with level, which also occurs. Thus changes in all three cases are in keeping with this idea of a blue "rod colour". Moreover, additivity deviations in B for test stimulus 530 nm, for 610 nm and for the combination of 478 and 546 nm, also show changes in a direction consistent with this idea. This is further supported by the fact that observer ERF could always restore to subjective acceptability trichromatic matches which had become mismatches due to change of luminance level, by adjusting just the B control, and always in the sense predicted.

Could tetrachromatic colorimetry show additivity failure?

Clarke (1973) has considered the possibility of additivity failure in a tetrachromatic system arising from a rather different cause to any previously discussed. Just as a system of three types of cone with spectral response curves of invariant *shape* must lead to an additive trichromatic system, four receptor types with such curves must give an additive tetrachromatic system. Tetrachromatic colour matches, where quantum absorption is balanced for all four receptor types, can only show additivity failure if the *shape* of one or more spectral absorption curves alters with changed conditions. Considering the spectral absorption curve for the rod photopigment rhodopsin, is the concentration high enough for bleaching at practical levels of illumination to change the concentration and hence the shape of the spectral absorption curve? Change of shape of the spectral absorption curve of rhodopsin with concentration is shown in Fig. 9 and it will be seen that for peak densities up to about 0·2, change of shape with concentration is not significant. Dartnall (1962) considers that the densities of rhodopsin in *human* rods are not high enough for significant change of shape to occur. Hence tetrachromatic additivity failure should not occur due to this cause.

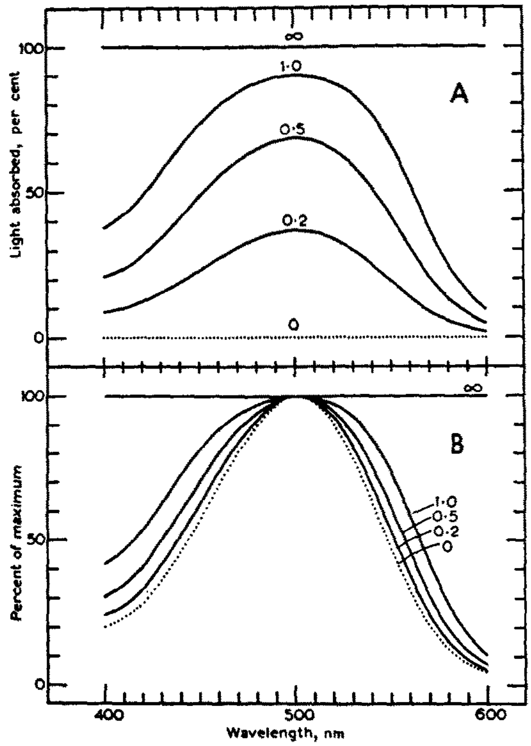

Fig. 9. Spectral absorption curves of rhodospin for various concentrations, after Dartnall (1962). The upper section A shows the absolute spectral absorption factor for different concentrations (indicated by the peak density figures shown). The lower section B shows the corresponding relative spectral absorption factors which determine the shape of the spectral responsivity function of the scotopic mechanism.

However Clarke (1973) pointed out that as extra foveal cones probably behave similarly to foveal cones in most characteristics, then one might expect tetrachromatic matches to break down at the very high luminance levels where foveal trichromatic matches are known to break down (Wright, 1936; Brindley, 1953).

Applications of the additive property of the tetrachromatic colour match

The additive property of the tetrachromatic colour match can be utilised to provide general systems of both colorimetry and photometry applicable at *all* luminance levels, i.e. scotopic, mesopic and photopic. Furthermore this can be done without making any assumptions concerning the four mechanisms (three types of cones, and rods) involved: i.e. each can be considered to act quite independently of the other three, or there may be interaction.

In view of the widespread computer facilities available today, the complexity of a calculation is unimportant as long as industry can be provided with valid systems of general colorimetry and photometry. A computer program to be developed will require a user in industry to feed in as data an absolute spectral power distribution: the print-out will then be given as (a) a set of three tristimulus values to specify the colour and (b) a single value to specify the "general trolands"—a term yet to be defined. It is essential that the *absolute* spectral power distribution should be fed in since both tristimulus values and "general trolands" are dependent on level. Item (a) will involve defining three suitable reference stimuli: because the trichromatic system is, in general, non-additive, no subsequent linear transformation is permitted. For item (b), "general trolands" must be defined in terms of a single reference stimulus, to be chosen to have a subjective brightness/power characteristic of continuous slope and is most likely to be a specified red: also to be defined is the method of photometry, probably to be direct heterochromatic brightness matching.

The computer program will provide two distinct kinds of permanent data, 1 and 2. Data 1 will allow the tetrastimulus values to be computed from the absolute spectral power distribution: these are defined similarly to conventional tristimulus values. It is this part of the computation that utilizes the additive property of the tetrachromatic system. Data 2 will allow in (a) computation of the tristimulus values and in (b) of "general trolands". It is important to realise that, in general, neither tristimulus values nor "general trolands" can be added or combined in any other way except via the tetrastimulus values.

The author's research programme is designed to determine Data 1 and Data 2 using facilities at N.P.L. initially and at City University later. Considerable work has already been done on 1, the tetrachromatic colour matching functions having been measured for one observer (Trezona, 1973b). Tetrachromatic colour matching functions together with a certain absolute spectral power distribution will allow tetrastimulus values (K, L, M, N) to be computed.

Data 2 will then provide the following functions, where R, G, B and A are derived from K, L, M, N: R, G and B are tristimulus values of stimuli $[R]$, $[G]$, $[B]$ and A is the value in terms of $[A]$, the photometric reference stimulus.

(a) $$\begin{cases} R = F_1(K, L, M, N) \text{ in specified units} \\ G = F_2(K, L, M, N) \text{ in specified units} \\ B = F_3(K, L, M, N) \text{ in specified units} \end{cases}$$

(b) $$A = F_4(K, L, M, N) \text{ “general trolands”.}$$

In order to determine these functions a certain combination of (K, L, M, N) is to be taken in one half of a colorimeter bipartite field. In the other half for (a) quantities of $[R]$, $[G]$, $[B]$ are adjusted for a colour match or for (b) the quantity of $[A]$ is adjusted for a brightness match. This must be done for a comprehensive lattice of combinations of (K, L, M, N) so as to obtain either the above functions or data for interpolation. This investigation may well prove to be long and laborious since restricting ourselves to n levels of which the lowest is the scotopic (and initially avoiding negative tetrastimulus values) will require n^4 separate determinations. It may well be that some “short-cuts” will be found, but these should only be used as the result of direct experimental evidence and not arising from unproved assumptions.

In the author's opinion it seems that, using experimental evidence of tetrachromatic additivity and trichromatic non-additivity and without making any physiological assumptions, this is the simplest possible sound approach that could be devised for general systems of colorimetry and photometry.

Acknowledgements—I am greatly indebted to Dr. F. J. J. Clarke for many helpful contributions. Thanks are due to Mr. P. G. Moss, Mr. E. R. Fulton and Dr. R. D. Lozano for acting as observers.

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.

Blottiau F. (1947) Les défauts d'additivité de la colorimétrie trichromatique. *Revue Opt. théor. instrum.* **26**, 193–201.

Bongard M. M., Smirnov M. S. and Friedrich L. (1958) The four-dimensional colour space of the extrafoveal retinal area of the human eye. *Visual Problems of Colour I*, pp. 325–330. NPL Symposium No. 8. H.M.S.O., London.

Brindley G. S. (1953) The effects on colour vision of adaptation to very bright lights. *J. Physiol., Lond.* **122**, 332–350.

Clarke F. J. J. (1960) Extrafoveal colour metrics. *Optica Acta* **7**, 355–384.

Clarke F. J. J. (1963) Further studies of extrafoveal colour metrics. *Optica Acta* **10**, 257–284.

Clarke F. J. J. (1973) Needs and prospects for a tetrachromatic system of large field colorimetry. *A.I.C. Second Congress: Colour 73*, pp. 319–324. Adam Hilger, London.

Clarke F. J. J. and Trezona P. W. (to be published).

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

Dartnall, H. J. A. (1962) The photobiology of visual processes. In *The Eye* (edited by H. Davson), Vol. 2, pp. 68. Academic Press, New York.

Lozano R. D. and Palmer D. A. (1967) The additivity of large-field colour matching functions. *Vision Res.* **7**, 929–937.

Lozano R. D. and Palmer D. A. (1968) Large-field colour matching and adaptation. *J. opt. Soc. Am.* **58**, 1653–1656.

Nimeroff I. (1964a) Colorimetry in parafoveal fields—II: Additivity failure. *J. opt. Soc. Am.* **54**, 833–838.

Nimeroff I. (1964b) Field trials of the 1959 C.I.E. Supplementary Standard Observer proposal. *J. opt. Soc. Am.* **54**, 696–704.

Palmer D. A. (1970) Colour matches which include equality of scotopic luminance. *Proc. 1st A.I.C. Congress, Stockholm*, 1969, pp. 166–171.

Palmer D. A. (1972) Tetrachromatic matches. *J. opt. Soc. Am.* **62**, 828–30.

Speranskaya N. I. (1958) Methods of determination of the co-ordinates of spectrum colours. *Visual Problems of Colour*—1, pp. 319–323. NPL Symposium No. 8 HM Stationery Office. London.

Speranskaya N. I. (1959) Determination of spectrum colour co-ordinates for twenty-seven normal observers. *Optics Spectrosc.* **7**, 424–428.

Stiles W. S. (1955) 18th Thomas Young Oration: The basic data of colour-matching. *Phys. Soc. Year Book*, pp. 44–65.

Stiles W. S. (1958) The average colour matching functions for a large matching field. *Visual Problems of Colour*—1, pp. 211–247 NPL Symposium No. 8. H.M.S.O., London.

Stiles W. S. (1963) NPL colour-matching investigation; addendum on additivity. *Optica Acta* **10**, 229–232.

Stiles W. S. and Wyszecki G. (1962) Field trials of colour-mixture functions. *J. opt. Soc. Am.* **52**, 58–75.

Trezona P. W. (1953) Additivity of colour equations. *Proc. Phys. Soc., Lond.* **B66**, 548–556.

Trezona P. W. (1954) Additivity of colour equations—II: *Proc. Phys. Soc., Lond.* **B67**, 513–522.

Trezona P. W. (1970) Rod participation in the “blue” mechanism and its effect on colour matching. *Vision Res.* **10**, 317–332.

Trezona P. W. (1972) A new method of large field colour matching leading to a more additive metric. *Proc. Helmholtz Memorial Symp. Colour Metrics*, Driebergen, September, 1971, pp. 36–49. A.I.C., Soesterberg, Holland.

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

Trezona P. W. (1973b) Tetrachromatic colour measurement. *A.I.C. Second Congress: Colour 73.*, pp. 319–324. Adam Hilger, London.

Wright H. and Wyszecki G. (1960) Field trial of 10° colour-mixture functions. *J. opt. Soc. Am.* **50**, 647–650.

Wright W. D. (1936) The breakdown of a colour match with high intensities of adaptation. *J. Physiol., Lond.* **87**, 23–33.

Wyszecki G. (1972) Discussion of paper by Trezona (1972) *Proc. Helmholtz Memorial Symp. Colour Metrics*. Driebergen, September 1971, p. 48. A.I.C., Soesterberg, Holland.

Résumé—Le système trichromatique de mesure des couleurs, s'il convient aux champs d'un diamètre de 2° environ, ne donne pas une métrique linéaire avec des champs plus grands à cause de la présence de bâtonnets. Pour en tenir compte il faut une égalisation colorée tétrachromatique et on peut donc espérer qu'un système tétrachromatique sera mieux linéaire. Le but de cet article est de comparer les additivités tétrachromatique et trichromatique pour divers stimulus tests. Un cas spécial de principe d'additivité, le métamérisme de luminance, constitue un test beaucoup plus sévère que le cas général: ce dernier n'est qu'altéré tandis que l'écart se concentre dans le premier. La plupart des expériences sont des mesures, mais certaines utilisent un jugement subjectif. Jusqu'ici on n'a pas pu déceler d'écart à l'additivité tétrachromatique, tandis que la non additivité trichromatique est très prononcée. Dans les cas examinés jusqu'ici les écarts à l'additivité trichromatique sont toujours dans un sens qui s'accorde avec l'hypothèse que la "couleur de bâtonnet" est bleue. On indique la voie dans laquelle l'additivité dans les égalisations tétrachromatiques de couleurs peut servir à développer des systèmes de colorimétrie et de photométrie, valides à tous les niveaux de luminance.

Zusammenfassung—Das trichromatische Maßsystem, das für Felder mit einer Ausdehnung von *ca.* 2° Gültigkeit hat, besitzt keine lineare Metrik für grössere Felder wegen des Vorhandenseins von Stäbchenrezeptoren. Diese werden im tetrachromatischen System berücksichtigt, daher kann man erwarten, dass dieses System linear ist. Dies zu untersuchen ist Aufgabe dieser Arbeit, indem tetrachromatische und trichromatische Additivität für eine Vielzahl von Testreizen verglichen wird. Ein Sonderfall des Additivitätsprinzips, die Leuchtdichtemetamerie, wird sorgfältiger getestet als der allgemeine Fall. Während der letztgenannte nur erwähnt wird, konzentriert sich die Untersuchung auf den erstgenannten Fall. In den meisten Experimenten wird gemessen, einige Male wird auch die subjektive Abschätzung untersucht. Bis jetzt wurde keine Abweichung von der tetrachromatischen Additivität entdeckt, obwohl die trichromatische Nicht-Additivität offensichtlich ist. Die Abweichung von der trichromatischen Addivität, soweit sie gemessen wurde, läuft immer in der Richtung, dass sie mit der Hypothese konsistent ist, dass die Farbe der "Stäbchen" blau ist. Es wird auf eine Möglichkeit hingewiesen, wonach die Additivitätseigenschaft des tetrachromatischen Systems dazu verwendet werden kann, ein allgemeines System der Colorimetrie und Photometrie zu entwickeln, das auf allen Leuchtdichteniveaus anwendbar ist.

Резюме—рихроматическая система измерения, хотя и подходит для полей размерами около 2°, не дает линейной метрики с большими полями, из за присутствия палочковых рецептороз. Их наличие учитывается в тетрахроматическом цветовом уравнении и, таким образом, можно ожидать, что тетрахроматическая система может быть линейной. Цель настоящей работы—испытать зто, путем сравнения тетрахроматической и трихроматической аддитивности при изменении тестовых стимулов. Специальный случай принципа аддитивности, яркостная метамерия, представляет много более серьезное испытание, чем общий случай: поскольку последний только касается его, первый на нем концентрируется. Большинство экспериментов включают измерение, но некоторые используют субъективное суждение. Все же не было никакого отклонения обнаружено в тетрахроматической адди9ивности, хотя трихроматическая неаддитивность была выражена. Так исслетуемая трихроматическая аддитивность всегда давала ошибку в направлении соответствушем гипотезе о том, что "пелочковый цвет"—Синий. Даны указания путей, которыми могут быть использованы аддитивные свойства трихроматического цветового уравнения для развития общей системы колориметрии и фотометрии, применимой при всех уровнях яркости.