Published in final edited form as:

J Opt Soc Am A Opt Image Sci Vis. 2014 April 1; 31(4): A186–A188.

Comparison of the Richmond HRR 4th edition and Farnsworth–Munsell 100 Hue Test for quantitative assessment of tritan color deficiencies

Katharina G. Foote, Maureen Neitz, and Jay Neitz*

Department of Ophthalmology, University of Washington, 750 Republican St., Seattle, Washington 98109, USA

Abstract

Drugs and environmental factors can induce tritan deficiencies. The Farnsworth–Munsell (FM) 100 Hue Test has become the gold standard in measuring these acquired defects. However, the test is time consuming, and color discrimination is confounded by concentration and patience. Here, we describe a test that compares six tritan plates from the HRR Pseudoisochromatic Plates 4th edition to 16 FM 100 Hue tritan caps. CIE Standard Illuminant C was reduced over five light intensities to simulate the effects of acquired losses in the S-cone pathway. Both tests showed quantitative differences in error rates with all light levels; thus they could serve equally well for assessing acquired deficiencies. However, compared to the FM 100, the HRR took subjects about 20–40 s per trial, making it more practical.

1. INTRODUCTION

The Farnsworth–Munsell (FM) 100 Hue Test is designed to evaluate and rank color discrimination [1]. It has the ability to distinguish people with normal color vision into classes of superior, average, and low discrimination abilities, and in people who have congenital color vision deficiencies, it gives an indication of the axis of confusion as along a blue–yellow or red–green axis or a generalized color vision loss [2]. The test was not originally intended for use of assessing acquired color vision deficiencies; however, it has become a standard for such assessment. The FM 100 is limited by factors such as the time to administer the test, as well intersubject variability due to concentration, patience, and cooperation [3].

The HRR Pseudoisochromatic Plates 4th edition has been shown to be an effective color vision test [4–6], and when compared to the FM 100, the HRR takes less time and is easier for the subject. The goal of the present study is to determine whether the HRR could serve as an effective alternative for quantitative assessment of acquired tritan-like defects. Although congenital tritan deficiencies caused by mutations in the short (S) wavelength sensitive cone opsin gene are rare, it has been shown that drugs and environmental factors

can induce tritan-like deficiencies [7] at a rate higher than any other type of acquired of color defect [8,9].

2. METHODS

A. Subjects

Ten color normal subjects as assessed by the Nagel Anomalo-scope, the HRR 4th edition, and the desaturated D15 volunteered to be tested on both the FM 100 caps and the HRR plates under low illumination to induce tritan deficiencies [10–14]. The subjects were five males and five females between the ages of 26 and 56.

B. FM 100

Only caps/plates that fall along the tritan confusion lines (Fig. 1) were used—from the FM 100, the first seven caps of box one (P, 85, 1, 2, 3, 4, 5), and caps 45–53 of box three. The caps were chosen using the FM 100 analysis of Smith *et al.* [15] and confirmed by spectroradiometry (SpectroCAL MKII Spectroradiometer, Cambridge Research Systems Ltd., 80 Riverside Estate, Sir Thomas Longley Road, Rochester, Kent ME2 4BH, UK). Subjects were instructed to place the caps in order of hue, and the first cap for each set was provided. The caps were scored using the official FM 100 scoring system found at <http://www.torok.info/colorvision/fm100.htm>. The caps not used were automatically scored as correct in order to merely isolate the tritan scores.

C. HRR

From the HRR, 18 flashcards with one figure per card (Fig. 2) were used. Of the total cards, six were tritan, six were tetartan, and six were “symbol-blank.” The cards were presented in one of four randomized orders. Subjects were instructed to identify the shape they saw on each card. They were told that some were “symbol-blank;” thus, guessing was not encouraged. Only the tritan cards were scored: one point for correctly identifying the shape, 0 points for an incorrect response, and –0.25 points if the subject identified a shape on a “symbol-blank” card.

D. Light Source

The light source used was a xenon arc lamp (OL 490 Agile Light Source, Gooch & Housego PLC, Dowlsh Ford, Ilminster, TA19 0PF, UK); a wide-band light source used to produce combinations of spectra at the light output. The light source was calibrated to match the chromaticity of CIE Standard Illuminant C. At reduced light levels, chromatic sensitivity for both blue/yellow and red/green stimuli is reduced. Here we were interested in blue/yellow deficiencies and in order to simulate the effects of acquired losses in the S-cone pathway, a series of five reduced illumination levels were determined empirically, spanning from the brightest where most subjects with normal color vision made no tritan errors to the dimmest, where responses were likely random. With the OL 490 Agile Light source it was possible to reduce the light level while exactly maintaining the chromaticity of CIE standard Illuminant C. Luminance values for the light levels were as follows (in cd/m^2): 1964.05, 188.78, 141.86, 48.75, and 22.95. Each subject was tested three times, usually spaced over several days, and the order of light levels was randomized for each trial. For each trial, the subject

was first tested on the 100 Hue, and then the HRR plates for each light level, before progressing to the next randomized light level.

E. Scoring

Averages for each subject for each test per light level were calculated. Next, the averages were divided by the total subjects' maximum error for each test in order to compute each subject's error score for each of the two tests. This scaling method was used because it conveniently made all scores fall between 0 and 1. These respective error scores were then plotted and two tailed t tests were performed for each point. This is shown in Fig. 3, and individualized scores are plotted in Fig. 4.

3. RESULTS

Both tests showed quantitative differences in error across light levels. The difference between the brightest and dimmest light levels was significant with a $p < 0.001$ for both the HRR and the FM100 Hue Test. The two tests were very similar in their ability to detect differences in color vision across the different light levels. A one-way ANOVA with correction for multiple comparisons showed no significant differences between the error scores on the HRR and the FM 100. However, subjects were much slower with the FM 100. The FM 100 required between 2 and 8 min for the subjects to complete only the selected tritan caps. The time to complete each trial varied depending on the light level and was cognitively stressful for the subjects. In comparison, HRR trials took only 20–40 s to complete.

4. CONCLUSIONS

Although the 100 Hue has become widely used in testing acquired color vision deficiencies, subjects reported the test to be time consuming and require much concentration, patience, and cooperation. In contrast, the HRR imposed no such discomfort or impairment and took less time to administer.

In addition to being speedy, for the HRR, subjects very rarely failed to detect symbols at the brightest light level (1964.0 cd/m^2). Even though none of the differences between the HRR and the 100 hue were statistically significant, from differences seen at the brightest light level, we would suggest that separating normal individuals from those with very subtle tritan color vision loss could be easier with the HRR than from the 100 Hue test. This is because we predict that almost all normals will make zero errors on the HRR while many normals would make a few mistakes on the 100 Hue.

Both tests showed quantitative differences in error rates with light level over the entire range and thus could serve for assessing the degree of acquired deficiencies; however, compared to the time-consuming FM 100, the HRR took the subjects about 20–40 s per trial for each condition, making it more practical.

Both the HRR and FM 100 tests were designed to be used by people with normal visual acuity. Acuity may be reduced in individuals with acquired color vision deficiencies and should be taken into account when choosing the appropriate test. The HRR might present

more difficulty than the FM 100 for those with low acuity. According to Birch, the HRR plates should not be used when visual acuity is below 6/60 m (~20/200 feet) [9].

Despite the FM 100 Hue Test becoming widely used in testing acquired color vision deficiencies, it is not ideal for such use. The HRR provides a rapid, accurate measure of color vision. The next step in comparing the two tests would be to administer the tests to patients with acquired tritan deficiencies.

Acknowledgments

This study was supported by NEI grants R01EY09303, P30EY01730, and RPB.

References

1. Farnsworth D. The Farnsworth-Munsell 100-hue and dichotomous tests for color vision. *J Opt Soc Am.* 1943; 33:568–578.
2. Birch J. Use of the Farnsworth-Munsell 100-Hue Test in the examination of congenital color-vision defects. *Ophthalmic Physiolog Opt.* 1989; 9:156–162.
3. Breton ME, Fletcher DE, Krupin T. Influence of serial practice on Farnsworth-Munsell 100-hue scores: the learning effect. *Appl Opt.* 1988; 27:1038–1044. [PubMed: 20531516]
4. Birch J, Platts CE. Colour vision screening in children: an evaluation of three pseudoisochromatic tests. *Ophthalmic Physiolog Opt.* 1993; 13:344–349.
5. Bailey JE, Neitz M, Tait DM, Neitz J. Evaluation of an updated HRR color vision test. *Vis Neurosci.* 2004; 21:431–436. [PubMed: 15518225]
6. Cole BL, Lian KY, Lakkis C. The new Richmond HRR pseudisochromatic test of colour vision is better than the Ishihara test. *Clin Exp Optom.* 2006; 89:73–80. [PubMed: 16494609]
7. Lawrenson JG, Kelly C, Lawrenson AL, Birch J. Acquired colour vision deficiency in patients receiving digoxin maintenance therapy. *Brit J Ophthalmol.* 2002; 86:1259–1261. [PubMed: 12386084]
8. Birch, J.; Chisholm, IA.; Kinnear, P.; Marre, M.; Pinckers, AJLG.; Pokorny, J.; Smith, VC.; Verriest, G. Acquired color vision defects. In: Pokorny, J.; Smith, VC.; Verriest, G.; Pinkcers, AJLG., editors. *Congenital and Acquired Color Vision Defects*. Grune & Stratton; 1979.
9. Birch, J. *Diagnosis of Defective Colour Vision*. Oxford University; 1993.
10. Verriest G. Further studies on acquired deficiency of color discrimination. *J Opt Soc Am.* 1963; 53:185–197. [PubMed: 13996879]
11. Boyce PR, Simons RH. Hue discrimination and light sources. *Lighting Res Technol.* 1977; 9:125–140.
12. Bowman KJ, Cole BL. A recommendation for illumination of the Farnsworth-Munsell 100-hue test. *Am J Optom Physiolog Opt.* 1980; 57:839–843.
13. Smith VC, Pokorny J, Pass AS. Color-axis determination on the Farnsworth-Munsell 100-hue test. *Am J Ophthalmol.* 1985; 100:176–182. [PubMed: 3874549]
14. Knoblauch K, Saunders F, Kusuda M, Hynes R, Podgor M, Higgins KE, Demonasterio FM. Age and illuminance effects in the Farnsworth-Munsell 100-hue test. *Appl Opt.* 1987; 26:1441–1448. [PubMed: 20454340]
15. Smith VC, Pokorny J, Yeh T. Pigment tests evaluated by a model of chromatic discrimination. *J Opt Soc Am A.* 1993; 10:1773–1784.

Fig. 1.

Two tritan confusion lines are drawn on the 1976 CIE diagram. The circles represent the 16 FM 100 caps presented to the subjects in this study measured empirically with a spectroradiometer.

Fig. 2.

18 plates selected from the HRR Pseudoisochromatic Plates 4th edition. There are six “symbol-blanks,” six tritan, and six tetartan. Tritan plates received a score of +1 or –1 points for correct or incorrect, respectively, “symbol-blank” plates received –0.25 for a false positive, and tetartan plates served as controls.

Fig. 3.

Average results of the HRR and FM 100 across 10 subjects. As light levels decrease, more errors are made. The five light levels were selected empirically based on where color normal observers made no errors, to where they made many errors.

Fig. 4.

Percent error scores for individual subjects for each of the five light levels. Top, HRR results; bottom, 100 Hue results. Observers tend to be generally similar in their ability to perform both tests. The points on the graphs are means from three sessions per observer.