Journal of Ophthalmology & Visual Sciences

Open Access

Austin Publishing Group

Review Article

Review of the Color Vision Tests Currently in Use

Almustanyir A*
Department of Optometry, King Saud University, Saudi Arabia

***Corresponding author:** Almustanyir Ali, Department of Optometry, King Saud University, Saudi Arabia

Received: April 28, 2020; **Accepted:** May 18, 2020; **Published:** May 25, 2020

Abstract

Purpose: Color vision testing is essential for people who perform tasks where color is used to convey information and accurate color judgments are essential for safe and efficient performance. A large number of color vision tests are currently available to screen for color vision deficiencies or detect the type and severity of the defect. Computerized color vision tests are now becoming more common in the clinical setting. Different programs are available that screen for color vision defects or perform both screening and diagnosis of the severity of the defect. This study aimed to review some of the color vision tests currently available in the market

Method: This study primarily focused on reviewing seven color vision tests: Ishihara, Hardy-Rand-Ritter plates (HRR), Waggoner PIP, Color Assessment and Diagnosis (CAD) test, Cone Contrast Sensitivity test, Farnsworth D15 (F-D15), and Waggoner D15 (W-D15).

Results and Conclusion: The majority of these tests showed very good agreement with the anomaloscope in assessments of a red-green color vision defect. The level of agreement, sensitivity, and specificity for most tests were comparable to the anomaloscope in terms of screening for color vision deficiency.

Keywords: Color vision test; Ishihara; HRR; Waggoner color vision test; CAD test; Cone Contrast Sensitivity test; Farnsworth D15; Waggoner

Introduction

The number of commercially available color vision tests in the market has increased, and it has become essential to determine which tests can be used in a convenient, valid, and reliable manner. A test can be considered good if it has the ability to accurately and quickly categorize subjects into Color Vision Defect (CVD) and Color Vision Normal (CVN) groups. However, a working knowledge of the test protocol is insufficient, and an understanding of the test design is essential to ensure high confidence in interpreting the results [1]. Moreover, it is necessary to understand the need and aim of color vision testing, which can be summarized as follows:

1. The test can screen and detect the presence of color vision deficiency. This situation would be the most common color activity.

2. The test has the ability to determine the type and severity of the color vision deficiency.

3. The test can assess the significance of color vision deficiency in those carrying particular color-dependent tasks. This commonly pertains to congenital red-green color vision deficiency rather than congenital tritan or acquired deficiencies. This group of tests would mimic the aspects of the intended occupation or a particular test at the actual place using the real system of work.

CVD can be classified into three general categories based on the number of primary colors that are required to make color matches. In monochromatic CVDs, patients require only one primary color to match all colored lights. Individuals with this form of CVD lack most of the normal cones and show a severe reduction in visual acuity. Because they have additional visual problems, they will not be discussed in this paper. The remaining categories of CVD psresent with a normal visual function and just CVDs, or they may show color vision independent of any visual problem. Dichromatism requires two primaries to match colors, whereas anomalous trichromats requires three primaries to make a match, but the amounts are significantly different from CVN.

Red-green color defects can be further classified based on whether the M-cone or L-cone photopigment is missing or different from the CVN population. Protanopes lack the long-wavelength sensitive pigment (L-cone), and deuteranopes are missing the medium wavelength sensitive pigment (M-cone). The anomalous trichromats can also be divided into parallel categories. Protanomalous trichromats possess an anomalous photopigment in their L-cone, while deuteranomalous trichromats have an anomalous photopigment in their M-cone [2].

Blue-yellow defects are very rare. Blue-yellow defects include the dichromatic (tritanope) and anomalous trichromat (tritanomalous) forms. Both tritanope and tritanomalous forms show problems with the S-cones. In tritanopes, the S-cone pigment is non-functioning, whereas in tritanomalous cases, the S-cones are only partially functional [3].

CVDs can be further categorized as congenital or acquired [1]. In the congenital form, the congenital visual system is otherwise normal except for the loss of color discrimination and the defect remains stable throughout, whereas in the acquired form, the CVD is related to an ocular disease or disorder and some other aspect of visual function is also affected by the condition. The defect can progress and regress along with the underlying condition [1,2]. Acquired

Citation: Almustanyir A. Review of the Color Vision Tests Currently in Use. J Ophthalmol & Vis Sci. 2020; 5(1): 1036.

Table 1: Scoring criteria for the color vision tests.

Test	Failure Criteria	Classification Criteria for Red-Green Defects	Severity Criteria
Ishihara (38 plates)	More than two errors on plates 1-17	Most of the errors or less distinct figures (on each plate) on the diagnostic plates in protan or deutan columns	NA
HRR 4th edition	**Blue-Yellow:** on screening plates (any error) **Red-Green:** >1 error on screening plates; and no errors on the diagnostic plates	The least number of errors in protan or deutan columns	**Red-Green** **Very mild**: Any red-green error on the screening plates and no errors on plates 11-20 **Mild**: Any error on plates 11-15 and no errors on the rest of the diagnostic plates. **Moderate**: Any error on plates 11-18 and no errors on plates 19 and 20. **Severe**: Any error on plates 19-20 **Blue-Yellow** **Mild:** The blue-yellow screening plates (any error) and plates 21-24 (no errors) **Moderate**: Plates 21-22 (any error) and no errors on plates 23 and 24. **Severe**: Any error on plates 23-24
Waggoner PIP	More than four errors on the red-green plates More than two errors on the blue-yellow plates.	More errors on protan vs. deutan diagnostic plates	**Red-Green** (defect with the majority of errors on the type of defect) **Mild**:more than 4, but less than 17 errors **Moderate:** more than 16, but less than 28 errors **Severe:** more than28 errors **Blue-Yellow** **Mild:** more than 3, but less than 7 **Moderate:** more than 6, but less than 9 **Severe:** more than 9
RCCT	Sensitivity value less than 75 for any cone type	Classification based on the minimum cone sensitivity	**Based on sensitivity value** **Mild:** more than 54, but less than 75 **Moderate:** More than 39, but less than 55 **Severe:** less than 40
CAD	**Screening mode**: percentage correct less than 66.6% for any color **Threshold mode:** Red-green SNU more than 1.77 OR Blue-yellow SNU more than 1.75	Protan, deutan, or unclassified based on the directions that have the highest thresholds.	Threshold values
F-D15	**Visual Inspection:** More than one major crossing **Color Difference Vector value:** C-index ≥1.78	**Visual Inspection:** depending on the error pattern to score sheet **Color Difference Vector angle size:** **Deutan** -20°< Angle ≤ -3° **Protan** -3°< Angle < 20° **Scotopic** 20°≤ Angle < 60° **Tritan** 60°≤ Angle ≤ 120° OR -120°≤ Angle ≤= -60°	NA
W-D15	**Color Difference Vector:** C-index ≥1.78	**Color Difference Vector angle size:** **Deutan** -20°< Angle ≤ -3° **Protan** -3°< Angle < 20° **Scotopic** 20°≤ Angle < 60° **Tritan** 60°≤ Angle ≤ 120° OR -120°≤ Angle ≤= -60°	NA

CVDs are less common in the general population. However, these defects are very common in the elderly. This is expected since the incidence of visual disorders also increases with age [2]. There are three types of acquired CVDs. Type I acquired red-green defects occur in photoreceptor/retinal pigment epithelium diseases. Patients with this type of defect tend to have protan defects. Type II acquired red-green defects are related to optic nerve diseases such as optic atrophies and optic neuritis [2]. Patients with this type of defect tend to have deutan defects. Type III acquired blue-yellow defects are the most common type of acquired CVD. This type of defect is defined by discrimination losses along the blue-yellow axis and is observed in macular degeneration, glaucoma, diabetes, nuclear cataract, and optic nerve disorders [2].

An accurate measurement of color discrimination is essential for jobs that require precise color vision to ensure safe and efficient

performance of the tasks, since patients with CVDs are at a higher risk of making errors in such tasks. A large number of color vision tests are currently available to detect CVD and estimate the patient's ability to discriminate colors. Computerized color vision tests are widespread and have become more common in the clinical setting. Some programs can screen for CVD or have the capability to screen and diagnose the severity of the defect. The purpose of this paper is to review the color vision tests that are commercially available in the market and widely used in the clinic. This study will focus on seven tests: Ishihara, Hardy-Rand-Rider plates (HRR), Waggoner PIP, Color Assessment and Diagnosis (CAD) test, Cone Contrast Sensitivity test, Farnsworth D15 (F-D15), and Waggoner D15 (W-D15).

Review of the Color Vision Tests

Ishihara

Ishihara Pseudo Iso Chromatic (PIC) plates are the most widely used color vision test to detect CVD. This test was first published in 1906 and was the first PIC test in commercial production. The test has been reprinted in many editions over the years and worldwide [4]. The 38-plate edition is considered to be the gold standard for red-green color vision screening [5]. The 38 plates version contains 25 numeral plates of one color embedded in a background of a different color and 13 pathway designed plates [5,6]. The numeral plates are divided into demonstration (plate number1), transformation (plate number 2-9), vanishing (plate number 10-17), hidden digit (plate number18-21), and classification (plate number22-25) [6]. Table 1 shows the criterion for this test.

Various studies on the efficiency of the Ishihara test have shown that the tester mains the "gold standard" for rapid identification of congenital red-green deficiencies. According to a study by Birch, the mean number of errors for all transformation and vanishing plates is around a maximum of 16 [6]. She found that the sensitivity and specificity for failing the test with a score of 4 or more errors are 98.7% and 94.1%, respectively [6]. She also found that it was challenging to use hidden digit plates. The reason for that is some CVN participants can read them, and the ability to read them depends on the subject's age. For example, 40% of the CVN participants aged between 20 to 30 yr can see the hidden Figure (6). The sensitivity of these plates was less than 50%. With respect to the classification plates, she found that 81.9% of protanopes and 93.4% of deuteranopes were correctly classified, whereas 18.1% of protanopes and 3.2% of the deuteranopes could not see either figure. In terms of anomalous trichromats, 46.6% of the protanomalous and 57.2% of the deuteranomalous trichromats were correctly classified while 6.7% of the protanomalous and 2.1% of the deuteranomalous trichromats could not see the classification numerals [6]. However, 40% of the protanomalous and 37.5% deuteranomalous trichromats were correctly classified in terms of luminance contrast (they could see both figures, but one was more distinct than the other). Overall, in terms of classification for missing one of the two figures or identifying one figure as more distinct than the other, 83.2% of protans and 94.1% of deutans were classified correctly [6]. Thus, the classification plates in the Ishihara test are precise in their ability to classify red-green defects. Hovis and Almustanyir [7] calculated the agreement between the Ishihara test and the anomaloscope using the AC1 coefficient of agreement value [8,9]. The reported an AC1 value of 0.86.They also found that the sensitivity and specificity were 0.95 and 0.91, respectively [7].

Hardy-Rand-Ritter plates

The HRR pseudo iso chromatic test was developed by Hardy, Rand, and Ritter in 1955 [10]. It is designed to detect tritan as well as red-green CVDs and grade their severity. The fourth edition was well designed as its colors aligned on the confusion lines and was isomeric with the original version of the test [4]. The fourth edition contains 24 plates that present with either one- or two-colored symbols (X, Δ, and O) within a gray background. These 24 plates include 14 plates designed to identify the type of deficiency (10 plates for deutans and protans and four plates for tritans) and grade the severity as mild, medium, or strong. For red-green defects, a defect is classified as severe when the observer makes one or more errors in the two plates with the most saturated colors, medium when the observer makes errors in the next three most saturated colors, and mild when the patient makes error with the least five saturated colors. Table 1 shows the criteria for this test.

Hovis and Almustanyir [7] compared the HRR and anomaloscope results for 65 CVD and 60 CVN participants. The found that the AC1 coefficient of agreement, sensitivity, and specificity for the red-green plates were 0.88, 0.95 and 0.91 respectively [7]. Another study conducted by Cole et al. reported that the average number of errors madefor red-green defects on the HRR screening plates was 4.97(out of six symbols), but the subjects did not make any errors on the tritan plates [11]. They found that the sensitivity and specificity were 1 and 0.96, respectively, when using two or more errors on the screening test figures as the failure criterion relative to the anomaloscope findings [11]. Eighty-six percent of their subjects were classified correctly as protan or deutan. This was similar to the percentages reported by Birch for the Ishihara. They also found that 31% of the CVDs were graded as showing a mild defect, 43% as showing a medium defect, and 26% as showing a severe defect [11]. The HRR results need to be compared with a standard measure of the severity to determine the validity of the severity grading. The Farnsworth D15 (F-D15) and anomaloscope tests are acceptable measures of severity. The F-D15 test can grade color defective subjects into two groups "pass" and "fail."Cole et al. found that, with one exception, subjects who were classified as mild by HRR passed the F-D15 test. Among those who were classified as medium by HRR, 40% failed the F-D15 test, and 85% of those classified as strong by the HRR failed the F-D15 [11].

In comparisons with the anomaloscope, Cole et al.found that the average anomaloscope range increased with HRR severity grading [11]. Subjects who were graded as a mild on the HRR test had an average range of 9.2 units on the anomaloscope (all of them had a range less than 30). This shows that they have good color discrimination, which agreed with the mild grade in HRR classification. Although the average range on the anomaloscope increased in the medium and strong grades, some of the subjects showing a small range in the anomaloscope were classified as medium or strong on the HRR. Twenty-four percent of the subjects classified as medium and strong had an anomaloscope range below 20 units, which is 90th percentile of the mild group. In addition, 37% of the dichromatic participants were classified as medium [11]. This shows that HRR is a suitable test to separate mild from strong, but it cannot separate medium from

strong [11].

Waggoner PIP test

The Waggoner PIP test is a new computerized color vision test that screens for red-green and blue-yellow color vision deficiencies. The old version of the test was introduced by Konan Medical. This version of the test is no longer available, and a similar version is available from Waggoner Diagnostics (Rogers, AR) as part of the Waggoner Computerized Color Vision Test (WCCVT). The Waggoner PIP test can classify the type of color vision defect and determine the severity of the defect. The plates used in this test are similar in design to the Ishihara plates. The test has 25 plates that are used to screen for red-green defects. These are followed by 12 screening and classification plates for the tritan defect. When a total of 5 errors are made, the red-green screening test ends, and the program switches to the blue-yellow plates before starting the red-green diagnostic plates. The tritan plates are to screen for a blue-yellow defect and can also classify the severity of the defect on the basis of the presentation of the plates, which vary systematically in their saturation. Individuals who cannot tell the more saturated colors are classified as showing a severe defect. The red-green diagnostic plates are administered next. This option will be taken automatically if the individual failed the red-green screening plates. Half of the diagnostic plates are for the protan defect, and the other half are for the deutan defect. Similarly, the saturation of the diagnostic tritan plates changes systematically in order to classify the severity of the defect. Each plate is presented for 2 s within a white background. After the number disappears, the patient enters the number that they saw or the "N" option if they did not detect a figure using the keyboard or the touch screen. Table 1 shows the criterion for this test. The failure on the red-green screening plates is 5 errors and more than 2 errors on the blue-yellow screening plates. The program determines the type and severity of the red-green defect based on the maximum number of errors made on the deutan and protan diagnostic plates. The severity of the blue-yellow defect is also based on the total number of errors.

A study conducted by Almustanyir and Hovis [12] validated a prototype of this new test and found that the Kappa coefficient of agreement (k), sensitivity, and specificity of the red-green plates relative to the anomaloscope were 0.95, 0.96, and 0.99, respectively [12]. They also determined the repeatability of the test, and they found that the k value between the first and second visit was high, with a value of 0.98 [12]. Another study by Hovis and Almustanyir (2017) assessed this test using the program on a Microsoft Surface Pro device [7]. They found that the AC1 coefficient of agreement values of the test with the anomaloscope was 0.9 for the red-green plates and 0.8 for the blue-yellow plates. The sensitivity and specificity relative to the anomaloscope were greater than 0.9. They also reported that the repeatability of the test with an AC1 coefficient of agreement values of the first and second visit of 0.98 for the red-green plates and 0.94 for the blue-yellow plates [7].

Rabin Cone Contrast Sensitivity Test

The RCCT is a new computerized color vision test to determine the threshold for letter recognition (cone contrast sensitivity). The principle of the RCCT involves measurement of the discrimination thresholds as a Tri vector from a grey reference using colors that should be missed by each of the three types of color vision deficiencies. Thus, each color is selected so that only one cone type is modulated as the saturation of the letter changes. The RCCT measures the discrimination threshold for the S-cone (detect a tritan defect), M-cone (detect a deutan defect), and L-cone (detect a protan defect) [13].

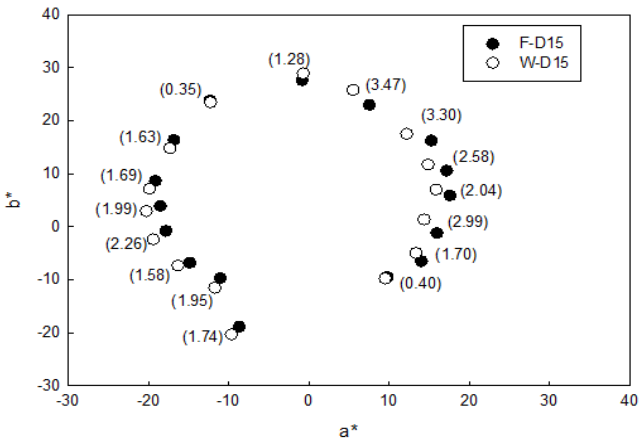

Figure 1: Chromaticity plot of a* and b* in normal space for the W-D15 (open circles) and F-D15 (closed circles) color vision test colors [25].

The test presents a single colored letter in the center of the computer's screen for approximately 5 s. The patient needs to respond by identifying the letter displayed on the screen. The order of testing is L-cone, M-cone, and S-cone contrast. Table 1 shows the criteria for this test. A sensitivity that is less than 75 for any cone type is considered to indicate failure in the test. The type of defect is based on the minimum cone sensitivity.

Two studies by Rabin in 2004 and 2011 validated this test [13,14]. The two studies reported that the RCCT had a perfect agreement with the anomaloscope with respect to whether the person had a CVN or a red-green CVD. That is, the specificity and sensitivity were both 1.0. Walsh et al. (2016) reported lower values with sensitivities of 0.97 for both right and left eye and specificities of 0.97 for the right eye and 0.96 for the left eye [15]. In a more recent study, Hovis and Almustanyir also determined the sensitivity and specificity of the test relative to the anomaloscope. They found that the sensitivity and specificity values were 0.95. They also calculated the AC1 coefficient of agreement, and reported an AC1 value of 0.9 [7].

Color assessment and diagnosis test

The CAD test is a color vision test that can screen for color vision deficiencies, measure chromatic discrimination around a gray reference, or both [16,17]. The CAD test assesses chromatic discrimination around a gray reference [17]. The CAD test consists of a gray-colored background and color stimulus that are buried in a background of dynamic luminance contrast noise. The background was made up of small individual squares, and the stimulus changes their luminance every 50 ms so that the screen looks as if it is scintillating. The colored stimulus moves in one of the four diagonal directions as the individual squares oscillate in luminances. An individuals' task is to respond by pressing a button to indicate in which direction the colored stimulus is moving. This test has two options. The first option is to screen for both red-green and blue-yellow defects. The

second option is to classify and quantify the severity of the defect. The observer's task is to detect the direction of the square. There are 4 possible diagonal directions. A four-alternative force-choice procedure is used to determine the observer's chromatic detection threshold in a specific direction within the CIE chromaticity chart [18]. The observer is instructed to maintain fixation on the center of the square and not to track the moving square in order to obtain the best results. If the observer is unsure of the direction, then s/he is encouraged to make the best guess since a response is necessary to continue the test. Different pass/fail criteria are allowed by this test, and this depends on the specific application of the results. Table 1 shows the recommended criteria for this test.

Barbur et al. (2006) evaluated the color vision of 250 individuals with a congenital red-green defect using CAD system [17]. The blue-yellow thresholds overlapped substantially with the CVN results, and these thresholds were not statistically significant. Nevertheless, there was a clear difference in the red-green threshold results. If ascore of >1.8 SNU is used to identify a patient with a red-green defect, the analysis in their Figure 3 suggests that both the specificity and sensitivity of the test were equal to a value of 1.0. Moreover, Seshadri et al. (2005) also reported specificity and sensitivity values over 90% for the WEB-based version [19]. Walsh et al. (2016) determined the specificity and sensitivity for the CAD threshold test, and they found that both values were near 0.85 [15]. Hovis and Almustanyir also reported that the AC1 coefficient of agreement, sensitivity, and specificity relative to the anomaloscope were 0.65, 0.92, and 0.60, respectively [7].

Farnsworth munsell D-15

The F-D15 color vision test was introduced to distinguish CVN and those with a mild CVD from individuals with a moderate-to-severe CVD [20]. The F-D15 is the current test that determines whether a soldier with a CVD has adequate color discrimination to perform his/her duties safely [7]. Subjects who failthe F-D15 are more likely to encounter problems in making color judgments in their everyday life or at work. Canadian military and civilian aviation authorities use the F-D15 test as a secondary test [7]. They test candidates with a screener test, and if the result is a failure, then the candidate must pass either the F-D15 or the Holmes–Wright Lantern Type A to qualify for an unrestricted pilot's license [7]. If the candidate passes the F-D15, they are considered as having met the color vision requirement to enter pilot training.

The patient's task is to arrange colored caps according to similarity by placing the color cap that is most similar to the previous cap set in the box. Major mistakes on the test are referred to as major crossings. The colors are approximately equally spaced around the hue circle, and the difference between adjacent colors is considered as an easily noticeable difference [20]. Errors on the F-D15 tests can be a transposition or a major crossing. The transposition occurs when a cap is placed only one to two positions from the correct arrangement, whereas major crossings occur when the caps from opposite positions of the hue circle are placed adjacent to one another. The test is commonly scored by visual inspection of the score sheet. Table 1 shows the recommended criterion for this test. Traditionally, two or more major crossings represents a failure.

The F-D15 can also be scored using Color Differences Vectors analyses [21]. The three parameters are specificity index (S-index), confusion index (C-index), and angular size. The S-index provides a measurement of how regularly the crossing is oriented. A high S-index value indicates a random arrangement. The C-index indicates the severity of the CVD. This index is correlated with the number of crossings and the total error score. The angular size provides a measurement of the type of CVD. The protan angles are larger than zero, and deutan angles are smaller than zero [21].

Fifty-five percent to 53% of participants with CVDs can pass the F-D15 test if one major crossing is allowed [22-24]. However, there is a possibility of a small number of dichromats (i.e., 3%) passing the F-D15 using this criterion. In the dichromat group, the percentage of those passing the test reduced to 1.5% when considering any crossing as a failure [24]. Hovis et al. reported the repeatability of the F-D15 using 116 red-green color defective subjects [23]. They found that if the failure criterion was two or more major crossings, the repeatability of the F-D15 was good with a kappa (κ) coefficient of 0.84 [23]. This value was less than the κ value of 0.96 calculated from Farnsworth's data. The reason for the difference in the κ values was that Farnsworth's population included a large number of CVNs, which would improve the repeatability of the test because CVNs rarely fail the test [23].

Waggoner CCVT D-15

The W-D15 program is a computerized version of the F-D15. The test is a part of the Waggoner CCVT test suite. It uses the same principle as the F-D15. The old version of the program requires the patient to drag the colored circle to the top of the display in order to use that color to "fill" one of the empty rectangles. The color selected should be the one that is most similar to the previous rectangle filled. The colors in the rectangles may be rearranged. The new version of the program presented colored circles on the left side of the screen. The patient is required to select a colored circle that is most similar to the previously filled circle and drag it to the first empty circle to complete a circle figure. The patient also has the option to rearrange the order. Table 1 shows the recommended criteria for this test using either the number of crossings or King-Smith indices.

Almustanyir and Hovis (2019) measured the chromaticity coordinates in a normal a* b* space for the F-D15 and W-D15. Figure 1 shows the results. The color difference between each cap of the two tests is listed in parentheses. The small color difference (ΔEs) between the two tests suggests that the tests should have the same level of difficulty [25].

Almustanyir et al. (2020) compared this test with the F-D15 test on 68 CVDs. They found that the AC1 agreement coefficient value was 0.88 [26]. They also calculated the Predictive value of Passing (PP) and the Predictive value of Failing (PF). The PR is the proportion of individuals who passed the W-D15 and F-D15 tests. The PF is the proportion of individuals who failed the W-D15 and F-D15 tests. They found that these two indices were good to excellent with PP and PF values of 0.87 and 1, respectively [26].

Discussion and Conclusion

Detection of color vision deficiency and determination of its severity has become more critical in many occupational environments. It is easy to use the number of errors that the observer

Table 2: Summary of all color vision test screening results reviewed in the study. The AC1 value, sensitivity, and specificity were calculated relative to the anomaloscope.

Test	Study	Sample size	Failure criteria	AC1 Coefficient of agreement	Sensitivity	Specificity
Ishihara (38 plates)	Birch (1997)[6]	CVD:401 N:100	>3 errors	-	0.98	0.94
	Hoivs & Almustanyir (2017) [7]	CVD:65 N:60	> 2 errors	0.86	0.95	0.91
HRR 4th edition	Cole *et al.* (2006)[11]	CVD:100 N:50	>1 error	-	1	0.96
	Hoivs & Almustanyir (2017) [7]	CVD:65 N:60	>1 error	0.88	0.95	0.91
Waggoner PIP	Almustanyir & Hovis (2015) [12]	CVD:47 N:75	>4 errors	0.95	0.96	0.99
	Hovis & Almustanyir (2017) [7]	CVD:65 N:60	>4 errors	0.9	0.9	0.9
RCCT	Rabin (2004)[13]	CVD:28 N:30	<75% sensitivity for any cone type	-	1	1
	Rabin et al. (2011)[14]	CVD:45 N:92	<75% sensitivity for any cone type	-	1	1
	Walsh, et al (2016)[15]	CVD:68 N:65	<75% sensitivity for any cone type	-	OD=0.97 OS=0.97	OD: 0.97 OS: 0.96
	Hoivs & Almustanyir (2017) [7]	CVD:47 N:75	<75% sensitivity for any cone type	0.9	0.95	0.95
CAD	Barbur, et al (2006)[17]	CVD: 250 N:238	>1.8 SNU	-	1	1
	Seshadri et al. (2005)[19]	CVD: 30 N:30	Any error in detecting the direction of the square	-	0.9	0.9
	Walsh et al. (2016)[15]	CVD:68 N:65	Any error in detecting the direction of the square	-	OD=0.86 OS=0.86	OD: 0.85 OS: 0.90
	Hoivs & Almustanyir (2017) [7]	CVD:47 N:75	>1.75 SNU	0.65	0.92	0.60

could make on a color screening test to separate CVNs from CVDs, but screening tests are not effective in predicting performance in occupational jobs such as aviation [18]. Several studies have reviewed the color vision tests used in clinical settings. The majority of the color vision tests evaluated in this study showed good agreement with the anomaloscope for the presence of red-green color vision defects in subjects. Table 2 summarizes the results of the tests reviewed in this study in terms of screening for red-green color vision deficiency (pass or fail the particular test) relative to the anomaloscope since it is the most common defect that could be seen in clinical settings [6,7,13-15,17,19,26]. The level of agreement, sensitivity, and specificity for most tests were comparable to the anomaloscope in terms of screening for color vision deficiency. The F-D15 and W-D15 tests were not included in the table since they are considered as alternative (secondary) tests that could be used to separate normal and mild CVDs from those with moderate to severe color vision deficiency. The printed tests (Ishihara and HRR are commonly used worldwide, and most clinicians use them in the clinic. These two tests are considered standard tests for screening for red-green color vision deficiency. The HRR has a slightly higher sensitivity in screening for red-green defects than the Ishihara, and it can also screen for blue-yellow defects. However, the printed tests have disadvantages. First, they are easy to memorize. Secondly, there is a high administrator bias. Third, the test cannot be integrating seamlessly with the electronic medical records. Fourth, the printed figures on the screening plates become relatively faint after prolonged usage, yielding a number of false positives. Fifth, the printed tests need to be administrated under specific lightning conditions. Finally, the test cannot monitor the workers' (for instance, pilots') color vision over their career since it cannot measure chromatic thresholds for both CVNs and CVDs. Nevertheless, the computer-based color vision test can eliminate all these disadvantages.

The computerized tests reviewed in this study can be used to detect minimum deficiencies, and some of them can assess the severity of discrimination losses by measuring the r-g and y-b thresholds. The results from these tests can provide an indication of the type of color losses. The CAD system measures chromatic thresholds relative to a gray background, and CCT can measure thresholds relative to other background hues. However, they use different stimuli configurations, size, and luminous noise in measuring thresholds so that a direct comparison between the three tests is difficult. Additional work is required to determine whether there is a simple scaling function that could be used to equate the thresholds measured with each system. Nevertheless, the computer tests need to calibrate the computer monitor regularly in order to ensure that the proper colors are displayed.

The typical occupational authority requirement of testing color vision accepts candidates who have a color vision defect and pass the F-D15. If the F-D15 is needed to be replaced by a newer test, then the new test should have very good agreement with the F-D15 and high predictive values for passing and failing the F-D15. The W-D15, which is designed to replace the F-D15, has excellent agreement with the F-D15. The PF value in the study by Hovis and Almustanyir (2017) was 100%, whereas the PP was slightly lower. This indicates that the W-D15 is slightly less sensitive than the F-D15. They reported that the reason for the lower PP is that the colored circles on the W-D15 are twice the size of the F-D15. Cole et al. (2006) reported that colors are slightly easier to identify for larger objects [27].

Assuming that the color vision test performance is essentially identical for a number of tests, then the other factors to consider in test selection are the cost and time to complete the test. In terms of the cost of the test, the CAD system is the most expensive, followed by the RCCT system and the Waggoner tests. However, CAD can measure color discrimination precisely. The printed test is the less expensive. The other factor is that a test that requires less time to administer may be preferred. The printed test would take less time to

complete than the computerized tests. CVN subjects could take less time on CAD and the Waggoner PIP test than CVD patients.

Acknowledgment

The author thanks the Deanship of Scientific Research and RSSU at King Saud University for their technical support. The author also would like to extend his appreciation to the College of Applied Medical Sciences Research Centre and the Deanship of Scientific Research at King Saud University for their funding of this research.

References

1. Lakowski R. Theory and practice of colour vision testing a review Br J Ind Med. 1969; 26: 173-89.
2. Pokorny J, Smith VC, Verriest G Congenital color defects. In Congenital and Acquired Color Vision Defects. 1979.
3. Deeb SS. Molecular genetics of color-vision deficiencies In Visual Neuroscience. 2004; 21: 191-96.
4. Dain SJ. Clinical colour vision tests Clinical and Experimental Optometry. 2004; 87: 276-93.
5. CIE. International recommendations for color vision requirements for transport. 2001; 143.
6. Birch J. Efficiency of the Ishihara test for identifying red-green colour deficiency Ophthalmic Physiol Opt. 1997; 17: 403-8.
7. Hovis J, Almustanyir A. Assessment of the Next Generation of Colour Vision Tests for Pilots and Aircrew. 2017; 9: 1-10.
8. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement Br J Math Stat Psychol. 2008; 61: 29-48.
9. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients. A study conducted with personality disorder samples BMC Med Res Methodol. 2013; 13: 1-61.
10. Walls GL. How good is the H-R-R test for color blindness? Optom Vis Sci. 1959; 36: 169-93.
11. Cole BL, Lian KY, Lakkis C. The new richmond HRR pseudoisochromatic test for colour vision is better than the ishihara test Clin Exp Optom. 2006; 89: 73-80.
12. Almustanyir A, Hovis JK. Military research ColorDx and printed color vision tests Aerosp Med Hum Perform. 2015; 86: 852-59.
13. Rabin J. Quantification of color vision with cone contrast sensitivity In Visual Neuroscience. 2004; 21: 483-85.
14. Rabin J, Gooch J, Ivan D. Rapid quantification of color vision: The cone contrast test. Investig Ophthalmol Vis Sci. 2011; 52: 816-20.
15. Walsh D V., Robinson J, Jurek GM, Capó-Aponte JE, Riggs DW, Temme LA. A performance comparison of color vision tests for military screening. Aerosp Med Hum Perform. 2016; 87: 382-87.
16. Barbur JL, Harlow AJ, Plant GT. Insights into the different exploits of colour in the visual cortex. Proc R Soc B Biol Sci. 1994.
17. Barbur J, Rodriguez-Carmona M, Harlow A. Establishing the statistical limits of 'normal' chromatic sensitivity. In CIE Expert Symposium, CIE Proceedings; 75 Years of the Standard Colorimetric Observer. 2006; 266: 613-46.
18. Rodriguez-Carmona M, O'Neill-Biba M, Barbur JL. Assessing the severity of color vision loss with implications for aviation and other occupational environments. Aviat Sp Environ Med. 2012; 83: 19-29.
19. Seshadri J, Christensen J, Lakshminarayanan V, Bassi CJ. Evaluation of the new web-based "Colour Assessment and Diagnosis" test Optom Vis Sci. 2005; 82: 10.
20. Farnsworth D. The Farnsworth Dichotomous Test for Color Blindness Panel D-15. New York Psychol Corp. 1947;
21. Vingrys AJ, King-Smith PE. A quantitative scoring technique for panel tests of color vision. Investig Ophthalmol Vis Sci. 1988; 29: 50-63.
22. Atchison DA, Bowman KJ, Vingrys AJ. Quantitative scoring methods for D15 panel tests in the diagnosis of congenital color vision deficiencies Optom Vis Sci. 1991; 68: 41-8.
23. Hovis JK, Ramaswamy S, Anderson M. Repeatability indices for the Farnsworth D-15 test. In: Visual Neuroscience. 2004; 21: 449-53.
24. Birch J. Pass rates for the Farnsworth D15 colour vision test. Ophthalmic Physiol Opt. Assessment of Current and Next Generation of Colour Vision Tests for Occupational Use. 2018; 58: 53-59.
25. Almustanyir , Hovis J. Trochromatic and dichromatic colorimetric analyses of the Farnsworth-Munsell D-15 color vision test. In: In Proceedings of the 29th Session of the CIE, CIE X046. Washington DC, CIE Cental Bureau Vienna. 2019.
26. Almustanyir A, Hovis J, Glaholt MG. Predicting the Farnsworth–Munsell D15 and Holmes–Wright-A lantern outcomes with computer-based color vision tests J Opt Soc Am A. 2020; 37: 1-10.
27. Cole BL, Lian KY, Sharpe K, Lakkis C. Categorical color naming of surface color codes by people with abnormal color vision Optom Vis Sci. 2006; 83: 879-86.