

Does the Farnsworth D-15 Match the HRR Test for Detecting Colour Vision Deficiency? A Comparison Against the Anomaloscope Reference Standard

Ali Almustanyir

aalmustanyir@ksu.edu.sa

King Saud University

Research Article

Keywords: colour vision deficiency, Hardy–Rand–Rittler test, Farnsworth D-15, anomaloscope, diagnostic agreement, Gwet's AC1

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11244557/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Purpose

The Hardy–Rand–Rittler (HRR) pseudoisochromatic plates and the Farnsworth D-15 are widely used to assess red-green colour vision deficiency (CVD), but have rarely been compared in the same individuals against the anomaloscope. This study examined how well the HRR classifies the type of defect, and how the HRR severity grade relates to the number of D-15 crossings and to the anomaloscope matching range.

Methods

One hundred and eight participants with red-green CVD confirmed by an Oculus HMC anomaloscope (68 deuterans and 40 protans; 55 dichromats and 53 anomalous trichromats) completed the HRR (4th edition) and the D-15 in a single session. Agreement between the HRR and anomaloscope classifications was assessed with Gwet's AC1 and associations with Spearman's rank correlation.

Results

The HRR screening plates were failed by 97.2% of participants, whereas 64.8% failed the D-15 (two or more crossings). The HRR protan/deutan classification agreed with the anomaloscope in 91.7% of participants (AC1 = 0.89, 95% CI 0.82–0.96). The number of D-15 crossings increased with the HRR severity grade ($\rho = 0.52$, $p < 0.001$). All 16 participants graded very mild or mild passed the D-15, but so did 33.3% of those graded medium and 7.3% of those graded severe. About half of the dichromats (27/55), mostly protanopes, were graded medium rather than severe. The anomaloscope matching range correlated moderately with the HRR severity grade ($\rho = 0.45$, $p < 0.001$) and strongly with the number of D-15 crossings ($\rho = 0.76$, $p < 0.001$).

Conclusions

The HRR has high screening sensitivity and classifies protan and deutan defects with almost perfect agreement with the anomaloscope. Individuals graded very mild or mild on the HRR can be expected to pass the D-15, but the medium and severe grades do not reliably separate dichromats from anomalous trichromats or predict the D-15 outcome.

Key Points

- People graded as very mild or mild on the Hardy–Rand–Rittler plate test all passed the Farnsworth D-15 arrangement test, so testing them with the second test may be unnecessary.

- The Hardy–Rand–Rittler plate test identified whether a red-green colour vision deficiency was of the protan or deutan type in agreement with the anomaloscope for about nine in ten people.
- About half of the people with dichromacy, the most severe red-green deficiency, were graded only as medium by the plate test, so its severity grades need cautious interpretation.

Introduction

Congenital colour vision deficiency (CVD) is a common inherited sensory disorder. It arises from X-linked mutations affecting the photopigment genes of the long- and middle-wavelength-sensitive cone photoreceptors, and affects approximately 8% of males and 0.5% of females of Northern European descent, with somewhat lower but still substantial prevalence reported across other populations [1]. CVD encompasses a spectrum of phenotypes, from mild anomalous trichromatism to complete dichromacy. The deficiency is the result of hybrid or deleted opsin genes within the L- and M-cone gene array on the X chromosome, with the specific genotype only partially predicting the severity of the resulting colour-discrimination loss [2].

Although congenital red-green CVD is not sight-threatening, it carries meaningful functional and occupational consequences. Affected individuals can experience difficulty with tasks requiring fine colour discrimination, and CVD can restrict entry into occupations such as aviation, the armed forces, electrical work, and certain branches of medicine and design, where accurate colour perception is safety-critical [3]. Beyond occupational restriction, CVD has been linked to everyday difficulties in food selection, clothing coordination, and interpretation of colour-coded information such as traffic signals, maps, and electronic displays, underscoring the clinical importance of accurate, efficient diagnostic testing [4, 5].

A range of clinical tools exists to detect and characterise CVD, spanning pseudoisochromatic plate tests, arrangement tests, and anomaloscopy, each differing in sensitivity, specificity, and the level of diagnostic detail provided [6]. Pseudoisochromatic plate tests and some arrangement tests are quick to administer and widely used for screening and severity grading in clinical practice [7]. Rayleigh matching on an anomaloscope, by contrast, remains the reference standard for diagnosing and classifying red-green CVD, as it directly probes the underlying photopigment matching function rather than relying on the perception of printed or manufactured colour stimuli, but its cost, the specialised equipment required, and the training needed to administer it limit its routine use to specialised clinical and research settings [8].

The Hardy-Rand-Rittler (HRR) test is a pseudoisochromatic plate test constructed on an opponent-colour model to detect and classify red-green and blue-yellow (tritan and tetartan) defects, and has been used clinically since its original development by the American Optical Company [9]. The first two editions had a relatively low specificity for screening red-green CVD, with more than 10% of normal trichromats missing at least one screening design [9–11]. Others have reported that the specificity and sensitivity of

the updated 4th edition are both high for adults, as is the accuracy of the protan/deutan classification [12, 13], but this has been questioned if the subjects are restricted to anomalous trichromats [11].

The Farnsworth D-15 (D-15) test is an arrangement test in which the examinee orders fifteen coloured caps by perceived hue similarity; the pattern of transpositions, or “crossings,” across the standard circular diagram indicates the presence, and to some extent the severity, of a colour-discrimination defect [14]. The test was designed to dichotomise individuals with CVD into those who are likely to have difficulty with everyday colour-related judgments and those who are not [14]. Several scoring approaches have since been proposed to quantify D-15 performance beyond a simple pass/fail outcome, translating the crossing pattern into a more continuous severity metric [15–17].

Despite widespread clinical use of both the HRR and the D-15, direct comparisons of their diagnostic performance against the anomaloscope reference standard, within the same sample and testing protocol, remain relatively limited for the updated HRR version. Previous studies reported that the HRR severity classification was moderately correlated with the anomaloscope range, and individuals who failed the D-15 were very unlikely to be classified as mild on the HRR [12, 13]. Establishing how these two commonly used clinical tools perform against the anomaloscope, and against each other, has direct relevance for choosing appropriate screening and diagnostic strategies in clinical optometry and ophthalmology practice, particularly in settings where anomaloscopy is not routinely available. The present study compared the diagnostic performance of the HRR and the D-15 against Rayleigh anomaloscope matching as the reference standard in a clinical sample of individuals with red-green congenital colour vision deficiencies. In particular, this study quantified the relationship between the number of D-15 crossings and HRR-graded severity and the agreement in protan/deutan classification between the HRR and the anomaloscope.

Methods

Participants

One hundred and eight participants with a red-green CVD took part. Table 1 lists the demographics. The participants were recruited from hospitals, schools, universities, and public places. Participants were included if they were aged 17–71 years and had best-corrected visual acuity of 6/9 or better in each eye. All participants were otherwise free of ocular abnormalities based on a short questionnaire to exclude individuals with known acquired CVD. All the participants reported being free from ocular or systemic diseases. Although the participants were permitted to use their regular distance or near-vision correction devices during testing, the use of tinted spectacles or contact lenses was not allowed. The study was approved by the University’s Institutional Review Board (E-22-6663) and adhered to the tenets of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to testing.

Colour Vision Assessment

All participants underwent colour vision assessment using three tests administered in a single session: the Rayleigh colour matching using the Heidelberg Multi-Color (HMC) Oculus anomaloscope (Oculus Optikgeräte GmbH, Wetzlar, Germany), the HRR (4th edition, Richmond Products) pseudoisochromatic plate test, and the D-15 arrangement test. The D-15 and HRR were performed binocularly, and the anomaloscope was performed monocularly using the participant's preferred eye. The D-15 and HRR were illuminated by a light-emitting diode (LED) lamp (NEW POWER, Model G45-14 E27) with a correlated colour temperature of 6500 K. The tests were placed on a white table with an average illuminance of 800 lx ($\pm 5\%$) in the horizontal plane. Test order was randomised across participants to minimise the influence of fatigue or learning effects on subsequent tests.

The HRR was administered according to the publisher's standard protocol instructions. Testing stopped if the participant passed the screening plates. For participants who failed the screening plates, the diagnostic plates were used to classify colour vision as protan- or deutan-axis deficiency, with severity further graded as very mild/mild, medium, or severe according to the number and type of diagnostic plates missed, per instructions [18].

The D-15 test was administered following the protocol outlined in the Appendix of Dain et al. [19]. Fifteen loose caps were randomly placed on a table in front of each participant, who was instructed to arrange them in order of perceived hue similarity, beginning from a fixed reference cap, and was permitted to rearrange the caps as required. Participants were encouraged to continue until all 15 caps were arranged in the box. Results were plotted on the standard circular diagram, and the number of diametrical crossings was recorded for each participant.

Data Analysis

Diagnostic outcomes from the HRR and D-15 were compared with anomaloscope classification and with each other. For the purposes of concordance analysis, anomaloscope subtypes were collapsed into a two-way classification (i.e. protan and deutan). Agreement between the HRR and anomaloscope axis classifications was quantified using raw percentage agreement and Gwet's AC1 statistic with 95% confidence intervals [20], interpreted using the Landis and Koch benchmarks [21]. The D-15 was only analysed for the number of crossings. Errors on the D-15 can vary from transpositions, in which a cap is placed only one or two positions from its correct place in the arrangement, to crossings, in which caps from opposite sides of the hue circle are placed adjacent to one another. Only crossings were counted, and two or more crossings were taken as a failure [22, 23].

The association between HRR severity grade (very mild/mild/medium/severe) and D-15 crossing count was assessed using Spearman's rank correlation coefficient and the Kruskal-Wallis test, given the non-normal distribution of crossing counts. Spearman's rank correlation was also used to assess the association of the anomaloscope matching range (in scale units) with the HRR severity grade and with the number of D-15 crossings. These analyses were performed using Python (version 3.11; pandas and

SciPy libraries). Gwet’s AC1 and its 95% confidence interval were calculated using the irrCAC library (version 0.4.4) for Python. A two-tailed p-value < 0.05 was considered statistically significant throughout.

Results

Sample Characteristics

Table 1 summarises the anomaloscope classification. All 108 participants were classified as having red-green CVD: 68 deuterans (63.0%) and 40 protans (37.0%).

Table 1
Sample characteristics and anomaloscope classification (n = 108). SD, standard deviation

Characteristic	Value
Total participants (n)	108
Age, years, mean ± SD (range)	28.8 ± 11.5 (17–71)
Sex, n (%)	Male: 98 (90.7%); Female: 10 (9.3%)
Deuteranomalous	40 (37.0%)
Deuteranope	28 (25.9%)
Protanope	27 (25.0%)
Protanomalous	13 (12.0%)

Pass/Fail Results

Because the D-15 is not a screening test [14], comparison with the HRR screening plates is of little value. This is further illustrated in that only 64.8% (n = 70) of the participants failed the D-15. In contrast, 97.2% (n = 105) failed the HRR screening plates, which is comparable to previous studies [12, 13].

Axis Classification: HRR versus Anomaloscope

Table 2 shows the agreement between the anomaloscope and HRR test with respect to the type of red-green defect. HRR axis assignment agreed with anomaloscope classification in 99 of 108 cases (91.7% raw agreement), with a Gwet's AC1 of 0.89 (95% CI 0.82–0.96), indicating almost perfect agreement. Of the 9 discordant cases, 1 was very mild, so no classification was possible and 3 were classified as normal by the HRR. All 4 were deuteranomalous with a matching range of 15 scale units or less. The remaining 5 involved axis misclassification between protan and deutan categories (3 deuterans classified as protan and 2 protans classified as deutan). If the comparison is restricted to only those individuals who were classified by the HRR, then the AC1 value increases slightly to 0.91 (95% CI 0.83–0.99; 95.2% raw agreement).

Table 2
Protan/deutan classification by the Hardy–Rand–Rittler (HRR) test compared with the anomaloscope (n = 108)

		Anomaloscope		
		Normal/Unclassified	Protan	Deutan
HRR classification	Normal/Unclassified	0	0	4
	Protan	0	38	3
	Deutan	0	2	61

Relationship Between HRR Severity Grade and D-15 Crossings

Figure 1 shows the relation between the HRR severity classification and the number of D-15 crossings. Table 3 shows the number of Farnsworth D-15 crossings according to HRR severity grade among participants who did not pass the HRR screening plates. Although there is substantial overlap in the medium and severe categories, there is a monotonic increase in the number of D-15 crossings with the HRR severity. A Spearman rank correlation confirmed a moderate, statistically significant positive association between HRR severity and D-15 crossings ($\rho = 0.52$, $n = 105$, $p < 0.001$) and the Kruskal-Wallis test indicated a significant overall difference in D-15 crossing count across the three HRR severity groups ($H = 31.62$, $df = 2$, $p < 0.001$). Notably, all 16 participants graded as very mild/mild by the HRR fell below the D-15 fail threshold of ≥ 2 crossings; however, 33.3% (16/48) of the medium and 7.3% (3/41) of the severe were also below the D-15 failure criterion. Of the 55 dichromats, 27 (49.1%) were graded as medium rather than severe; these were mostly protanopes (23 of the 27 protanopes compared with 4 of the 28 deuteranopes).

Table 3
Number of Farnsworth D-15 crossings by Hardy–Rand–Rittler (HRR) severity grade for the participants who failed the HRR screening plates (n = 105). SD, standard deviation

HRR Severity	n	D-15 crossings, mean $\pm$ SD	Range
Very Mild/Mild	16	0.1 $\pm$ 0.3	0–1
Medium	48	5.4 $\pm$ 4.6	0–13
Severe	41	7.9 $\pm$ 3.9	0–14

Relationship With the Anomaloscope Matching Range

The anomaloscope matching range was moderately correlated with the HRR severity grade ($\rho = 0.45$, $n = 105$, $p < 0.001$) and strongly correlated with the number of D-15 crossings ($\rho = 0.76$, $n = 108$, $p < 0.001$). Because the dichromats accepted all or nearly all of the range, the correlations were also calculated for the anomalous trichromats alone. In this group, the matching range was moderately correlated with both

the HRR severity grade ($p = 0.40$, $n = 50$, $p = 0.004$) and the number of D-15 crossings ($p = 0.43$, $n = 53$, $p = 0.001$).

Discussion

This study compared the diagnostic performance of the HRR pseudoisochromatic plate test and the D-15 arrangement test against anomaloscope classification in a sample of 108 individuals with red-green CVD. Three principal findings emerged: (1) as expected, the HRR demonstrated substantially higher sensitivity for detecting CVD than the D-15 (97.2% vs. 64.8%); (2) HRR axis classification (protan/deutan) showed almost perfect agreement with anomaloscope classification ($AC1 = 0.89$); and (3) D-15 crossing count correlated moderately with HRR severity grade with no individuals with a very mild/mild defect failing the D-15.

Detection Sensitivity

As mentioned in the Introduction, the D-15 is designed to dichotomise individuals with a CVD who have trouble making daily colour-related judgments from those who do not [14]. As such, it is not a screening test and would be expected to have a lower sensitivity. The D-15 failure rate in this study was slightly higher than reported by Birch [24], and likely reflects a bias towards more severe defects resulting from our recruitment targeting individuals with known or suspected CVD. Although the sensitivity value in this study is marginally lower than the 0.99 and 1.0 reported in previous studies [12, 13], our results confirm that the HRR has excellent screening sensitivity for red-green CVDs in adults. Based on our results, individuals with CVDs who are most likely to pass the HRR are deuterans with an anomaloscope matching range of 15 scale units or less. This result is consistent with our previous study that showed mild deuterans can pass the HRR and some may also pass other screening tests [13].

Interestingly, the 5.7% (3/53) of anomalous trichromats who passed the HRR is slightly less than the 7.2% reported by Birch [11]. Nevertheless, the values were similar. In her study, the majority of these individuals failed the Ishihara, indicating that the Ishihara is a better test for identifying individuals with a mild defect. Whether these individuals in this study would also pass the Ishihara test is uncertain.

Axis Classification Accuracy

HRR axis classification (protan/deutan) showed almost perfect agreement with anomaloscope classification (91.7% raw agreement, $AC1 = 0.89$). This finding aligns closely with previously reported HRR classification accuracies that varied between 86% and 100% [12, 13]. In Cole et al.'s study, 11% were unclassified and 3% were misclassified [12]. This study found more individuals were misclassified than unclassified, but the total misclassified and unclassified was less than 20% in both studies. The difference could be due to different recruiting protocols. Cole et al. recruited subjects through the Colour Vision Clinic and we used advertisements. The difference could have biased our sample towards more severe CVDs, which are less likely to be unclassified but may be more likely to be misclassified. Although we did not directly compare the HRR to the Ishihara test, given the similar classification results with

previous studies, our results provide indirect confirmation that the HRR is superior to the Ishihara in classifying the defect as protan or deutan [12, 13, 25]. Comparison with the D-15 was not performed because of the number of subjects who would be considered as unclassified because they passed the test.

Severity Gradient

Given the number of dichromats who were classified as medium, it is not surprising that the positive correlation between HRR severity grade and D-15 crossings was only moderate. Nevertheless, the results confirm Cole et al.'s findings [12] that individuals who are classified as mild on the HRR will pass the D-15, individuals with a severe defect on the HRR are more likely to fail the D-15, and close to 50% (27/55) of the dichromats will be classified as having a medium defect on the HRR. The results are also similar to other studies which measured saturation thresholds using the Colour Assessment and Diagnosis (CAD) test. Similar to our results, individuals who passed the D-15 had lower mean thresholds than those who failed the D-15, but the ranges of thresholds for the two groups had substantial overlap [26, 27]. Most of the dichromats classified as medium were protanopes (23 of the 27 protanopes compared with 4 of the 28 deuteranopes). The reason for this is uncertain, but it could be related to the slight drift away from the lines of confusion shown by Dain's colorimetric analysis [28] or a luminance artefact. Because the range of crossings for the medium and severe groups is similar, combining the test results provides limited information about the severity of the defect beyond the very mild/mild case. The result that there was only moderate agreement between the two tests is related to a number of factors. The first is that the D-15 is assessing hue discrimination for a relatively constant colour difference of colours from around the hue circle, whereas the HRR is estimating saturation discrimination for colours along the dichromatic lines of confusion. Thus, the D-15 is a "functional" colour vision assessment and the HRR is estimating thresholds, so a moderate relationship is a reasonable outcome. Also, both tests are designed to align with "average" dichromatic lines of confusion, so that a given individual with ocular media, receptor or post-receptor differences from the average would have a different outcome of one or both tests.

Clinical Implications

The first implication is that, in situations where the D-15 is used to determine occupational suitability and the HRR is used as the initial screening test, individuals who have been diagnosed as very mild/mild do not necessarily need to take the D-15, since they have an extremely high probability of passing the D-15. A very mild/mild defect on the HRR also strongly suggests that the person should not encounter many problems related to colour judgments in their daily lives. Although a small minority of individuals classified as severe on the HRR may still pass the D-15, this discrepancy should be considered suspicious given concerns about practice effects on the D-15 [29, 30]. In these situations, further testing may be necessary to resolve the disparity. However, passing the D-15 with a medium defect on the HRR is a possible outcome, as is a large number of crossings on the D-15.

Limitations

As mentioned, the sample was drawn from hospitals, schools, universities, and a few from public places and consisted entirely of participants with anomaloscope-confirmed CVD, which precluded specificity estimation and limits generalisability to unselected screening populations for the HRR, where the proportion of true normal trichromats would be higher. Whether an increase in the proportion of individuals with a mild CVD would change the correlation between the HRR severity classification and D-15 crossings is uncertain. Nevertheless, both tests show limitations in identifying dichromats.

Conclusion

As expected based on the test design, the HRR test demonstrated a screening sensitivity similar to previous studies. It also showed almost perfect axis-classification agreement with the anomaloscope and, because of its higher sensitivity, was capable of classifying more individuals than the Farnsworth D-15. Testing individuals who have a very mild/mild defect on the HRR with the Farnsworth D-15 may be unnecessary. The correlation between the HRR severity classification and the number of D-15 crossings helps interpret clinical findings when both tests are used.

Declarations

Author Contributions

Ali Almustanyir was responsible for the study conception and design, data collection, analysis and interpretation of the data, and writing of the manuscript, and read and approved the final manuscript.

Conflict of Interest

Funding: This research was funded by the Ongoing Research Funding Program (ORF-2026-1168), King Saud University, Riyadh, Saudi Arabia. The author has no relevant financial or non-financial interests to disclose.

Data Availability

The data from this study are available from the corresponding author upon reasonable request.

Ethics Approval

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Institutional Review Board of King Saud University (No. E-22-6663). Written informed consent was obtained from all participants.

Acknowledgments:

The author extends his appreciation to the Ongoing Research Funding Program (ORF-2026-1168), King Saud University, Riyadh, Saudi Arabia.

References

1. Neitz J, Neitz M. The genetics of normal and defective color vision. *Vision Res.* 2011;51:633–51. <https://doi.org/10.1016/j.visres.2010.12.002>
2. Deeb SS. Molecular genetics of colour vision deficiencies. *Clin Exp Optom.* 2004;87:224–9. <https://doi.org/10.1111/j.1444-0938.2004.tb05052.x>
3. Cole BL. The handicap of abnormal colour vision. *Clin Exp Optom.* 2004;87:258–75. <https://doi.org/10.1111/j.1444-0938.2004.tb05056.x>
4. Simunovic MP. Colour vision deficiency. *Eye.* 2010;24:747–55. <https://doi.org/10.1038/eye.2009.251>
5. Almustanyir A. A global perspective of color vision deficiency: awareness, diagnosis, and lived experiences. *Healthcare.* 2025;13:2031. <https://doi.org/10.3390/healthcare13162031>
6. Dain SJ. Clinical colour vision tests. *Clin Exp Optom.* 2004;87:276–93. <https://doi.org/10.1111/j.1444-0938.2004.tb05057.x>
7. Fanlo Zarazaga A, Gutiérrez Vázquez J, Pueyo Royo V. Review of the main colour vision clinical assessment tests. *Arch Soc Esp Oftalmol (Engl Ed).* 2019;94:25–32. <https://doi.org/10.1016/j.oftale.2018.08.010>
8. Cole BL. Assessment of inherited colour vision defects in clinical practice. *Clin Exp Optom.* 2007;90:157–75. <https://doi.org/10.1111/j.1444-0938.2007.00135.x>
9. Birch J. Clinical use of the American Optical Company (Hardy, Rand and Rittler) pseudoisochromatic plates for red-green colour deficiency. *Ophthalmic Physiol Opt.* 1997;17:248–54. <https://doi.org/10.1111/j.1475-1313.1997.tb00052.x>
10. Baraas RC. Poorer color discrimination by females when tested with pseudoisochromatic plates containing vanishing designs on neutral backgrounds. *Vis Neurosci.* 2008;25:501–5. <https://doi.org/10.1017/S0952523808080632>
11. Birch J. Identification of red–green colour deficiency: sensitivity of the Ishihara and American Optical Company (Hard, Rand and Rittler) pseudo-isochromatic plates to identify slight anomalous trichromatism. *Ophthalmic Physiol Opt.* 2010;30:667–71. <https://doi.org/10.1111/j.1475-1313.2010.00770.x>
12. 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. <https://doi.org/10.1111/j.1444-0938.2006.00015.x>
13. Almustanyir A, Hovis JK. Military research ColorDx and printed color vision tests. *Aerosp Med Hum Perform.* 2015;86:852–9. <https://doi.org/10.3357/AMHP.4326.2015>
14. Farnsworth D. The Farnsworth-Munsell 100-hue and dichotomous tests for color vision. *J Opt Soc Am.* 1943;33:568–78. <https://doi.org/10.1364/JOSA.33.000568>
15. Bowman KJ. A method for quantitative scoring of the Farnsworth Panel D-15. *Acta Ophthalmol.* 1982;60:907–16. <https://doi.org/10.1111/j.1755-3768.1982.tb00621.x>

16. Vingrys AJ, King-Smith PE. A quantitative scoring technique for panel tests of color vision. *Invest Ophthalmol Vis Sci.* 1988;29:50–63.
17. Foutch BK, Stringham JM, Lakshminarayanan V. A new quantitative technique for grading Farnsworth D-15 color panel tests. *J Mod Opt.* 2011;58:1755–63.
<https://doi.org/10.1080/09500340.2011.573881>
18. Bailey JE, Neitz M, Tait DM, Neitz J. Evaluation of an updated HRR color vision test. *Vis Neurosci.* 2004;21:431–6. <https://doi.org/10.1017/S0952523804213463>
19. Dain SJ, Atchison DA, Hovis JK. Limitations and precautions in the use of the Farnsworth-Munsell dichotomous D-15 test. *Optom Vis Sci.* 2019;96:695–705.
<https://doi.org/10.1097/OPX.0000000000001420>
20. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol.* 2008;61:29–48. <https://doi.org/10.1348/000711006X126600>
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–74. <https://doi.org/10.2307/2529310>
22. Almustanyir A, Alduhayan R, Alhassan M, Bokhary K, Alabdulkader B. Evaluation of the Waggoner computerized D15 color vision test using an iPad device. *J Opt Soc Am A.* 2021;38:1647–55.
<https://doi.org/10.1364/JOSAA.431987>
23. Almustanyir A, Alduhayan R, Alhassan M, Hovis JK. Clinical analysis of the Konan–Waggoner D15 color vision test using the Surface-Pro display. *J Opt Soc Am A.* 2023;40:A91–8.
<https://doi.org/10.1364/JOSAA.477206>
24. Birch J. Pass rates for the Farnsworth D15 colour vision test. *Ophthalmic Physiol Opt.* 2008;28:259–64. <https://doi.org/10.1111/j.1475-1313.2008.00566.x>
25. Birch J. Efficiency of the Ishihara test for identifying red-green colour deficiency. *Ophthalmic Physiol Opt.* 1997;17:403–8. <https://doi.org/10.1046/j.1475-1313.1997.97000227.x>
26. Evans BEW, Rodriguez-Carmona M, Barbur JL. Color vision assessment-1: visual signals that affect the results of the Farnsworth D-15 test. *Color Res Appl.* 2021;46:7–20.
<https://doi.org/10.1002/col.22596>
27. 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:A1–10.
<https://doi.org/10.1364/JOSAA.381305>
28. Dain SJ. Colorimetric analysis of four editions of the Hardy-Rand-Rittler pseudoisochromatic tests. *Vis Neurosci.* 2004;21:437–43. <https://doi.org/10.1017/S0952523804213475>
29. Ng JS, Morton WA. Case report: invalidation of the Farnsworth D15 test in dichromacy secondary to practice. *Optom Vis Sci.* 2018;95:272–4. <https://doi.org/10.1097/OPX.0000000000001184>
30. Ng JS, Liem SC. Can the Farnsworth D15 color vision test be defeated through practice? *Optom Vis Sci.* 2018;95:452–6. <https://doi.org/10.1097/OPX.0000000000001218>

Figures

