How accurate is an on-line test for colour vision deficiency?

Kari Bjerke Gjærde [1], Anne Kristin Kvitle [1,*], Phil Green [1] and Peter Nussbaum [1]

[1] Norwegian University of Science and Technology; kari.gjarde@ntnu.no; anne.kvitle@ntnu.no; phil.green@ntnu.no; peter.nussbaum@ntnu.no

***** Corresponding author: anne.kvitle@ntnu.no

Abstract

This study explores the accuracy and specificity of an on-line version of a standardized colour vision deficiency test – the Hardy-Rand-Rittler test (HRR) performed in an uncontrolled environment.
A group of 25 observers (18 with a colour vision deficiency and 7 with normal colour vision) that had previously been tested in a controlled setting participated, and the results from the on-line test was compared with previous results.
The on-line test successfully predicted the main results of the physical test of all 25 observers. The test also predicted the deutans with an accuracy of 92 %.

Keywords: *colour vision, colour vision deficiency (CVD), colour vision test*

INTRODUCTION

Background

Colour vision deficiency (CVD) is a common phenomenon, affecting as many as one in twelve males (8%) and one in 200 females (0.5 %) according to Birch (2012). Studies reported by Cole (2007) claim that as many as 20 to 30 percent of adults with abnormal colour vision are not aware of their condition. This suggests that colour vision is not tested as often as it should be.

Colour vision deficiency is detected and classified by visual tests, including pseudoisochromatic plates (PIP), sorting or arrangement tests, and colour matching tests. The latter includes the anomaloscope, which is considered the most accurate and the only existing test capable of reliably distinguishing anomalous trichromacy from dichromacy. One of the most widely used tests is the Ishihara PIP test, as proposed by Shinobu Ishihara in 1917 (Ishihara (1972)). Other PIP tests include the HRR test and the Waggoner tests (Waggoner 2021). Digital tests include the Cambridge Color Test (CCT) (Mollon and Regan 2000) and the Colour Assessment and Diagnosis (CAD) (Seshadri et al. 2005) test that are designed for display, in addition to the digital versions of the Waggoner tests (both PIP and sorting) as described and evaluated by Ng et al. (2015). According to Murphy (2015), digital tests are close to being the ideal colour vision test as they are easy to administer, quick to conduct and accurate, and often cheaper to produce and maintain. However, the digital tests are device dependent and a colour-calibrated display is required to guarantee accuracy of a physical test (Tsai et al. 2003), Ng et al. 2015). On-line tests include variations of the Ishihara and the Farnsworth D-15, but a review by Zarazaga et al. (2019) questions the scientific validity of published studies of such tests and Murphy (2015) suggest that a specific on-line version of the Farnsworth-Munsell 100-Hue is an acceptable alternative to prove excellent colour vision rather than a tool to diagnose CVD subjects.The Ishihara test has some limitations in determining the type and severity of CVD, which can also be caused by the examiner mis-interpreting the results. According to Cole (2007), common mistakes include using the wrong fail criterion and the assumption that a high number of errors indicate a severe CVD. Cole et al. (2006) find the Richmond HRR better than the Ishihara test, which is supported by Aldewachi (2013). Both studies conclude that HRR is as good as the Ishihara in detecting CVD, and that HRR has an advantage over Ishihara in grading the severity of the defect. Cole et al. (2006) suggest that the HRR test could be the test of choice for clinicians who wish to use a single test for colour vision.

The study presented in this paper explore to what extent a digital version of the HRR test conducted on-line in an uncontrolled setting could verify a colour vision deficiency on subjects with a confirmed and categorized condition. Specifically, asked:

- How do the results from the on-line test compare to the physical results?
  - Will the test detect a colour vision deficiency among the CVD subjects? (i.e. what is the sensitivity of the test)
  - Will the test confirm normal colour vision for colour normal observer (CNO)? (i.e. what is the specificity of the test)
- For persons with CVD: Is the outcome the same as the physical test with respect to deutan and protan deficiencies?

METHOD

The on-line test was designed as an uncontrolled web-based test, were the subjects performed the test at their choice of display (laptop display, external display mobile phone, external display) and environment. The chosen test was the HRR, where the plates of the physical HRR test (add reference) were scanned and prepared in Adobe Photoshop and exported for web at 600 px and embedded sRGBIEC61966-21 colour profile. The results were then compared with those for the same observers on the physical HRR test.

HRR test

According to the HRR test instructions (Good-Lite.com 2021), the HRR test is required to be conducted in CIE standard Illuminant C or a close approximation of it, where the examinator shows the HRR booklet about 30 inches from the eyes of the subject. The booklet contains 24 plates, of which the first four plates are demonstration plates to explain and train the subject. These four plates show the subjects what, where and how many symbols can be included. After the demonstration, there is a screening series of six plates, where the examinor asks the subject how many and what symbols they see, and where they are. If the subject sees all symbols, their color vision is considered to be normal and the test ends here. If there are any errors (i.e. symbols not seen) the test will continue based on which of the screening plates the subject made error(s) on. The number of visible symbols in each category is listed in table 1.

Number of symbols		
Plate 5 – 10 (Screening)	**Plate 11 – 20 (protan/deutan)**	**Plate 21 – 24 (Tritan)**
10	18	8

Table 1: Number of symbols in screening and diagnonostic plates.

The type of CVD is analyzed based on the count of the number of symbols that the subject detects, and the severity is determined according to the last group of plates in which the error occurs. If the subject make errors in the screening, but passes the rest of the test, the screening plates should be repeated with the booklet rotated 90 or 180 degrees to give the symbols a different location on the page.

Digital test

Typeform (Typeform 2021) was used as the on-line data acquisition platform, as it has a good adaptation for mobile devices and allows the display of a large image area of the plates.

The subjects received an individual ID code and a link to the test. Before performing the HRR test, demographic data (age and gender) and viewing conditions (display type, display size and illumination) were collected by self-reports from the subjects.

In order to explore the subjects' physical environment (size and quality of display and illumination), different grey scales were made to obtain feedback on the user's display settings (contrast, brightness etc.); one bright, one mid tone and one in the darker area. Adjustments were made in the L* lightness, with 2 levels difference in the three versions. Also, a grey scale with greater differences (and easier to separate and quantify) was used in the introduction, as a "training" part to the grey scales. An example of a grey scale is illustrated in Figure 1.

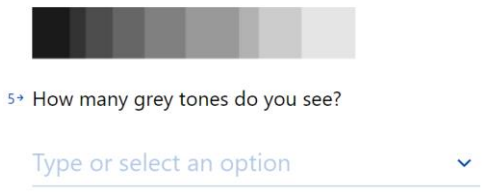
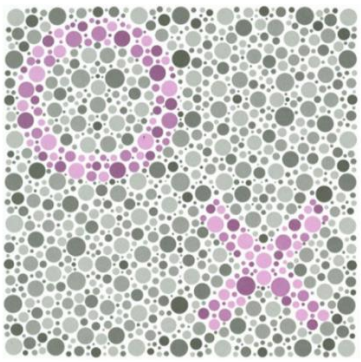


Figure 1: Grey scale.

HRR – conducting the test

With the physical HRR test, the examiner sits together with the patient, and explains the parts of the test, symbols on plates etc. For the digital version, a "translation" of the examiner's explanation was shortened and given as an introduction text, followed together with an illustration as shown in Figure 2. The first four plates of a HRR test are for demonstration, and the result on this part shows that almost everyone understood this and answered correctly on these four, non-scored, plates. In contrary to the original HRR test, all subjects scored all the plates in the digital test, regardless of the screening results.



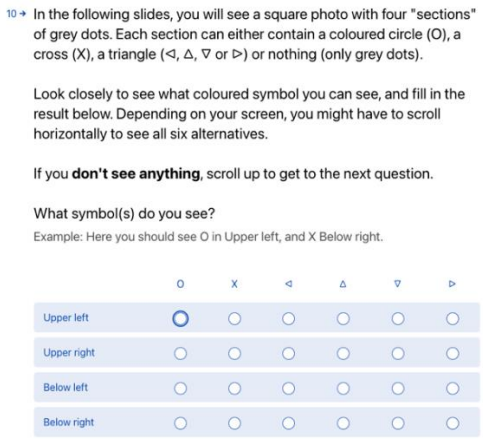


Figure 2: Screenshot of instructions before demonstration plates.

Observers

This experiment involved two groups of observers: a group of previously tested colour vision deficient observers (CVDs) and a reference group of people with previously tested normal colour vision (Colour Normal Observers (CNOs)). The CVD observers were recruited among participants from previous experiments and had previously been tested in a controlled setting (Kvitle et al. 2018), performing the HRR test followed by the Farnsworth D-15 and Lanthony D-15. The CNO observers were recruited from students at NTNU that had been tested during lab practice.

RESULTS AND DISCUSSION

A total of 25 observers conducted the uncontrolled on-line test, with 18 previously confirmed CVD (13 deutans and 5 protans) and 7 previously confirmed CNO by the physical HRR test. Their age range spans from twenty-two to 63. Of these 25, 18 had been found to have a colour vision deficiency by the test, and a further 7 participants were found to have normal colour vision.

How are the results compared to the physical results?

The following categorization is based on the previous test results from the physical tests in the lab. If a subject does not recognize one or more symbols in plate 5–10, the subject is categorized as a CVD and continues to the plates 11–24 (protan and deutan) or 21–24 (tritan) for diagnostic categorization and severeness (Good-Lite.com 2021). In this on-line test, all of the subjects assessed all plates.

CVDs

All observers previously categorized as CVD obtained results that categorized them as CVDs in the on-line test, as they did not detect symbols in the screening plates 5 - 10 and in the diagnostic plates 11 – 20. The average number of errors in each category is shown in table 2. For more details on the diagnostic series (plate 11–20), please see below regarding deutans and protans.

Average number of symbols not detected (CVD)		
Plate 5 – 10	**Plate 11 – 20**	**Plate 21 - 24**
5.50	7.61	0

Table 2: Average number of symbols not detected by CVD observer group in screening and diagnostic plates.

CNOs

All the CNOs clearly confirmed that they have normal colour vision. All observers saw all the symbols in plates 5 – 10 after two of them were asked to redo the test. In plates 11 – 20, three persons missed out on one symbol, which was corrected to one after a second test. The average number of symbols not detected is listed in table 3.

Average number of symbols not detected (CNO)					
First trial			**Second trial**		
Plate 5 – 10	**Plate 11 – 20**	**Plate 21 - 24**	**Plate 5 – 10**	**Plate 11 – 20**	**Plate 21 - 24**
0.50	0.50	0	0	0.17	0

Table 3: Average number of symbols not detected by CNO observers in two trials.

Given that the CNOs should see all the symbols all the time, this gives an uncertainty of 5.3 % the first time of testing, and 0.8 % after retesting.

Specificity and sensitivity

The results show that all the participants in the in each group scored according to expected results according to their performance on previous CV test. This gives the test 100 % sensitivity, with regards to the previous categorization of physical testing. The specificity of the test is also 100 %.

How accurate is the test regarding categorization of CVD?

All the CVDs were confirmed CVD in the digital test. Based on the HRR Score sheet, the on-line test distinguished 12 deutans, 1 protan and 5 uncertain results. Here, the term "uncertain" refers to an equal number of errors in each category (protan or deutan). The results from the on-line HRR test and the previous results are listed in table 4.

	Protan	**Deutan**	**Uncertain**
On-line HRR	1	12	5
Physical tests	5	13	0

Table 4 : Results on protan/deutan from on-line and physical test.

This indicates that the on-line test predicts a deutan with an accuracy of 92 %. Among the CVDs the overall uncertainty is 28 % when it comes to categorize the type of CVD. 5 subjects could not be clearly categorized, whereas 4 previously categorized as protans and one categorized as deutan. It has to be noted that the deutan subject is categorized as nearly normal CV based on previous tests.

Only five persons had previously tested as protan, so the results cannot be used to draw firm conclusions about the performance of the test. One of the reasons for the limited number of protan subjects are the frequency of this category in the general population, which is also reflected in similar studies in the literature. Based on the scores for diagnostics in plates 11 – 20, where the scores for protan and deutan can be eight in each category (8P – 8D), the confirmed deutans are more clear (1P–7D, 3P–8D, 2P–5D etc.), whereas the score for the confirmed protans are unclear where three of the uncertain scoring equal in each category (8Protan(P)–8Deutan(D), 2P–2D, 3P–3D). However, when examining the severity of the condition (mild, medium or strong) which is recognized by the last group of plates in which the errors occur, the errors for the protan group occurs in the protan category. This is an interesting finding, which should be further examined.

CONCLUSION AND FUTURE WORK

Our study showed that an uncontrolled on-line version of the acknowledged HRR test can detect and confirm both CVD and normal colour vision among the group of observers. This is promising, as such test can be implemented as a tool for CV screening test in uncontrolled psychophysical experiments where either a normal colour vision is required, or a CVD is requested. However, based on our results, the test does not appear able to clearly differentiate the categories of CVD which is often required in colour vision experiments involving CVD simulation or daltonization methods. To verify, or reject this, more observations are needed.

ACKNOWLEDGEMENT

The authors express our gratitude to the Good-Lite Company for giving permission to use the HRR test in this experiment. We would also thank our observers for their patience and time. This research was funded by the Research Council of Norway, through project no. 221073 “HyPerCept – Colour and quality in higher dimensions”.

REFERENCES

Aldewachi, A. 2013. Comparison of two colour vision tests used in current ophthalmic practice. *Tikrit Medical Journal* 19:255-268.

Birch, J. 2012. Worldwide prevalence of red-green color deficiency. *Journal of the Optical Society of America A* 29: 313-320.

Cole, B. L. 2007. Assessment of inherited colour vision defects in clinical practice. *Clinical & experimental optometry* 90:157–175.

Cole, B. L., Lian, K. Y., & Lakkis, C. 2006. The new Richmond HRR pseudoisochromatic test for colour vision is better than the Ishihara test. *Clinical & experimental optometry* 89: 73–80.

Dain, S.J., Clinical colour vision tests, *Clinical and Experimental Optometry* 87: 276-293.

Farnsworth, D. 1947. The Farnsworth Dichotomous Test for Color Blindness: Panel D-15, ed: Psychological Corporation.

Good-Lite.com. “HRR Test Guide”. Available at https://www.good-lite.com/collections/color-vision-tests/products/730005. Accessed July 30, 2021.

Hardy-Rand-Rittler (HRR) pseudoisochromatic test for colour vision, 4th Ed., Test plates and instruction book, Richmond Products

Ishihara, S. 1972. Tests for Colour-Blindness – 24 Plates (24 Plates ed.), Kanehara Shuppan Co., Ltd.

Lanthony, P. 1978. The desaturated panel D-15. *Documenta Ophtalmologica* 46: 185.

Mollon JD, Regan, BC. 2000. Cambridge colour test handbook. Cambridge: Cambridge Research Systems.

Ng, J.S., Self, E., Vanston, J.E., Nguyen, A. L., Crognale, M. A. 2015. Evaluation of the Waggoner Computerized Color Vision Test. *Optometry and Vision Science* 92:480-486.

Kvitle, A. K., H. Oddløkken, P. Green, and P. Nussbaum. 2018. Methods for psychophysical assessment of colour difference by observers with a colour vision deficiency. In: *2018 Colour and Visual Computing Symposium (CVCS), Proceedings of the Colour and Visual Symposium.*

Murphy, R.A. 2015. *Comparing color vision testing using the Farnsworth-Munsell 100-Hue, Ishihara compatible, and digital TCV software.* Master’s thesis. Pacific University Oregon College of Optometry.

Seshadri J., Christensen J.O., Lakshminarayanan V.F., Bassi C.J. 2005. Evaluation of the New Web-Based "Colour Assessment and Diagnosis" Test. *Optometry & Vision Science* 82:882-5.

Simunovic, M. 2010. Colour vision deficiency. *Eye* 24: 747-755.

Tsai, Y.-C., Chen, H.-S., Luo, M. 2016. Screen-based color vision test for digital quantification of deutan color defects using a color-calibrated monitor. *Journal of the Society of Information Display* 24:487 – 500.

Typeform. 2021. Available at www.typeform.com. Accessed July 30, 2021.

Zarazaga, A.F., Vásquez, A.J., and Royo, V.P. 2019. Review of the main colour vision clinical assessment tests. *Archivos de la Sociedad Española de Oftalmología* 94: 25-32.

Waggoner diagnostics. 2021. Available at https://waggonerdiagnostics.com/pages/information. Accessed July 30, 2021.