

# Normative values for a tablet computer-based application to assess chromatic contrast sensitivity

Lakshmi Bodduluri<sup>1</sup> · Mei Ying Boon<sup>1</sup> · Malcolm Ryan<sup>2</sup> · Stephen J. Dain<sup>1</sup>

Published online: 14 April 2017  
© Psychonomic Society, Inc. 2017

**Abstract** Tablet computer displays are amenable for the development of vision tests in a portable form. Assessing color vision using an easily accessible and portable test may help in the self-monitoring of vision-related changes in ocular/systemic conditions and assist in the early detection of disease processes. Tablet computer-based games were developed with different levels of gamification as a more portable option to assess chromatic contrast sensitivity. Game 1 was designed as a clinical version with no gaming elements. Game 2 was a gamified version of game 1 (added fun elements: feedback, scores, and sounds) and game 3 was a complete game with vision task nested within. The current study aimed to determine the normative values and evaluate repeatability of the tablet computer-based games in comparison with an established test, the Cambridge Colour Test (CCT) Trivector test. Normally sighted individuals [ $N = 100$ , median (range) age 19.0 years (18–56 years)] had their chromatic contrast sensitivity evaluated binocularly using the three games and the CCT. Games 1 and 2 and the CCT showed similar absolute thresholds and tolerance intervals, and game 3 had significantly lower values than games 1, 2, and the CCT, due to visual task differences. With the exception of game 3 for blue-yellow, the CCT and tablet computer-based games showed similar repeatability with comparable 95% limits of agreement. The custom-designed games are portable, rapid,

and may find application in routine clinical practice, especially for testing younger populations.

**Keywords** Color vision · Tablet computer · Games · CCT Trivector test · Repeatability · iPad

## Introduction

Chromatic contrast sensitivity (CCS) is defined as the ability to discriminate between stimuli based on their chromaticity difference alone, independent of any luminance contrast (Jacobs, 1993). Tests of CCS have clinical utility to assess or detect deficiencies in color vision. Color vision deficiencies may be congenital or secondary to disease and often manifest as a decreased ability to differentiate between shades of a color or between two or more colors. Ocular conditions such as diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma are known to affect color vision prior to affecting visual acuity (VA) (Greenstein, Hood, Ritch, Steinberger, & Carr, 1989; O'Neill-Biba, Sivaprasad, Rodriguez-Carmona, Wolf, & Barbur, 2010). In individuals with diabetes mellitus (DM), color vision impairment has been documented to emerge in the early stages of the disease and may precede the development of DR (Feitosa-Santana et al., 2006; Kurtenbach, Schiefer, Neu, & Zrenner, 1999; Ventura, Costa, et al. 2003). Therefore, the assessment of acquired color vision abnormalities is a useful measure for clinicians as it may be the earliest manifestation of a disease condition.

Color vision is usually assessed in the clinical environment using screening tools rather than estimating threshold or sensitivity, as threshold assessment traditionally requires specialized, calibrated laboratory equipment, which requires a trained person to administer the test. Currently available clinical tests of color vision include screening tests such as

✉ Lakshmi Bodduluri  
l.bodduluri@unsw.edu.au

<sup>1</sup> School of Optometry and Vision Science, The University of New South Wales, High St, Kensington, Sydney, NSW 2052, Australia

<sup>2</sup> Department of Computing, Macquarie University, Sydney, NSW 2109, Australia

pseudoisochromatic (PIC) plate tests and tests that determine the color discrimination thresholds. Screening tests are designed to fail individuals with even mild color vision deficiencies and have a single pass-or-fail criterion. Tests that measure color discrimination thresholds quantify the discrimination abilities of an individual and also help to understand the severity of color vision deficiency. Discrimination tests must be designed with care. Some discrimination tests have been found to require certain cognitive skills to perform the test (Cranwell, Pearce, Loveridge, & Hurlbert, 2015; Dain & Ling, 2009) or may have a theoretical bias towards tritan errors (Dain, 2004; Lakowski, 1969; Melamud, Hagstrom, & Traboulsi, 2004). Moreover, these tests require trained clinician administration and instrumentation. Along with these clinical color vision tests, there are other commercially available computer-based tests such as the Cambridge Colour Test (CCT) (Mollon & Regan, 2000), the modified CCT for children (Goulart et al., 2008) and the Colour Assessment and Diagnosis test (CAD) (Seshadri, Christensen, Lakshminarayanan, & Bassi, 2005) which do measure color thresholds and may be found in clinics that specialize in color vision. If color threshold measures are to be used as a diagnostic indicator of visual system dysfunction in chronic visual conditions such as DR, AMD, and glaucoma, then it would be beneficial if computer-based tests could be developed to have a simple interface that allows patients to self-administer the tests, and have a small form factor (Anderson, Burford, & Emmerton, 2016). Despite their diagnostic utility, the CCT and CAD tests are not designed for unsupervised self-administration and their physical dimensions preclude easy portability. Moreover, self-monitoring requires repeated testing, so it would also be beneficial if such tests were designed to be attractive to maintain the attention and compliance of users (Anderson et al., 2016).

Touchscreen technology found in tablet computers and personal mobile telephones have been harnessed as a supporting platform for the development of mobile health applications (apps). Technology has become increasingly personalized so that many individuals are in close proximity to such devices. Tablet computers are affordable, portable, and handy, and so are well placed to be developed into portable vision tests to monitor for any changes in vision due to systemic/ocular conditions. Several authors have reported the development and use of vision testing apps on the tablet computers (Aslam et al., 2013; Dorr, Lesmes, Lu, & Bex, 2013; Kollbaum, Jansen, Kollbaum, & Bullimore, 2014; Mulligan, 2013; Rodriguez-Vallejo, Remon, Monsoriu, & Furlan, 2015) and also as the tools for psychophysical experiments (Turpin, Lawson, & McKendrick, 2014). Thus, the ability to assess color vision routinely either in clinical practice or at home would be facilitated by the development of a portable and easy-to-administer instrument, such as a tablet computer-based app. The ability for self-administration of color vision tests would facilitate a monitoring role for such

technology, which may be addressed by the gamification of these vision tests. Furthermore, it has been suggested that presenting vision tests as computer games (Abramov et al., 1984) or on portable tablet computers as digital games (Nguyen, Do, Chia, Wang, & Duh, 2014) may help in the assessment of vision in an engaging manner. In fact, tests of color vision which are designed for their entertainment value have reached the popular news (MailOnline-Australia, 2015), indicating that color vision tests have scope to be made more fun and engaging.

Therefore, three tablet computer-based games were developed with different levels of gamification, to assess chromatic contrast sensitivity (to detect small departures from normal chromatic contrast sensitivity but not to diagnose any congenital color vision abnormalities). The purpose of the present study was to determine the normal range of chromatic contrast thresholds (tolerance intervals) using the custom-designed tablet computer-based app and to test the repeatability of these new designs in comparison with an established test, the CCT Trivector test.

## Methods

### Participants

A total of 100 healthy control participants [median (range) age of 19.0 years (18–56 years)] with a VA of 6/6 or better, measured with a Bailey Lovie LogMAR VA chart, were recruited into the study. The VA was measured both monocularly and binocularly with the participants' habitual correction. As monocular and binocular vision was better than 6/6 for all participants, the binocular values are reported here for brevity. All participants were screened for red–green or blue–yellow congenital color vision deficiencies using Ishihara's pseudoisochromatic plate test (Ishihara, 1917) and the Standard Pseudoisochromatic Plate I (Mäntyjärvi, 1987) tests, respectively. The study protocol was approved by the Human Research Ethics Advisory (HREA: #14225) of the University of New South Wales and all the procedures followed the tenets of the Declaration of Helsinki. All the participants gave their written informed consent, after explanation about the study procedures prior to any testing. All participants were tested for their chromatic contrast thresholds using the three custom-designed tablet computer-based games and the CCT Trivector test.

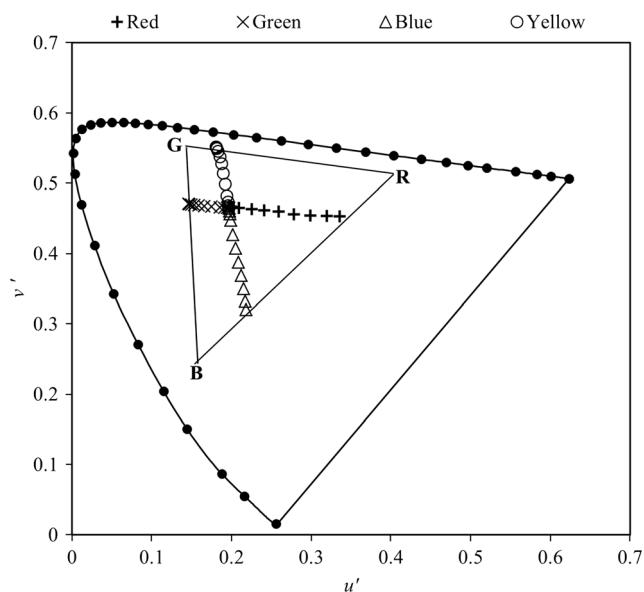
### Chromatic contrast sensitivity tests and procedure

#### *The tablet computer-based application*

The iPad mini retina display (Apple Inc.; display resolution: 2,048 × 1,536 pixels at 326 pixels per inch, 8-bit resolution and screen size: 7.9 in., with a measured screen luminance of

406 cd/m<sup>2</sup>) device was calibrated for its display characteristics prior to the development of the vision testing application (Bodduluri, Boon, & Dain, 2016). These characteristics informed the design of the visual stimuli used in the vision tests through the presentation of stimuli that were within the capabilities of the device's display to accurately produce. Three custom-designed games were developed with different levels of gaming elements and were designed to assess chromatic contrast sensitivity. The three games had the same stimulus background with a fixed chromaticity of  $u' = 0.197$ ,  $v' = 0.466$  (corresponding to a central gray: R = G = B = 127 with a luminance of 88 cd/m<sup>2</sup>) and a stimulus that varied in chromaticity relative to the background according to a psychophysical staircase procedure. Dithering was employed in the presentation of the stimuli to enable accurate chromaticity display. In the test design, games 1 and 2 employed luminance noise in both the background and the stimulus, in order to avoid any luminance clues to assist in the identification of the stimulus. Game 3 did not employ any luminance noise due to its specific design characteristics (explained below).

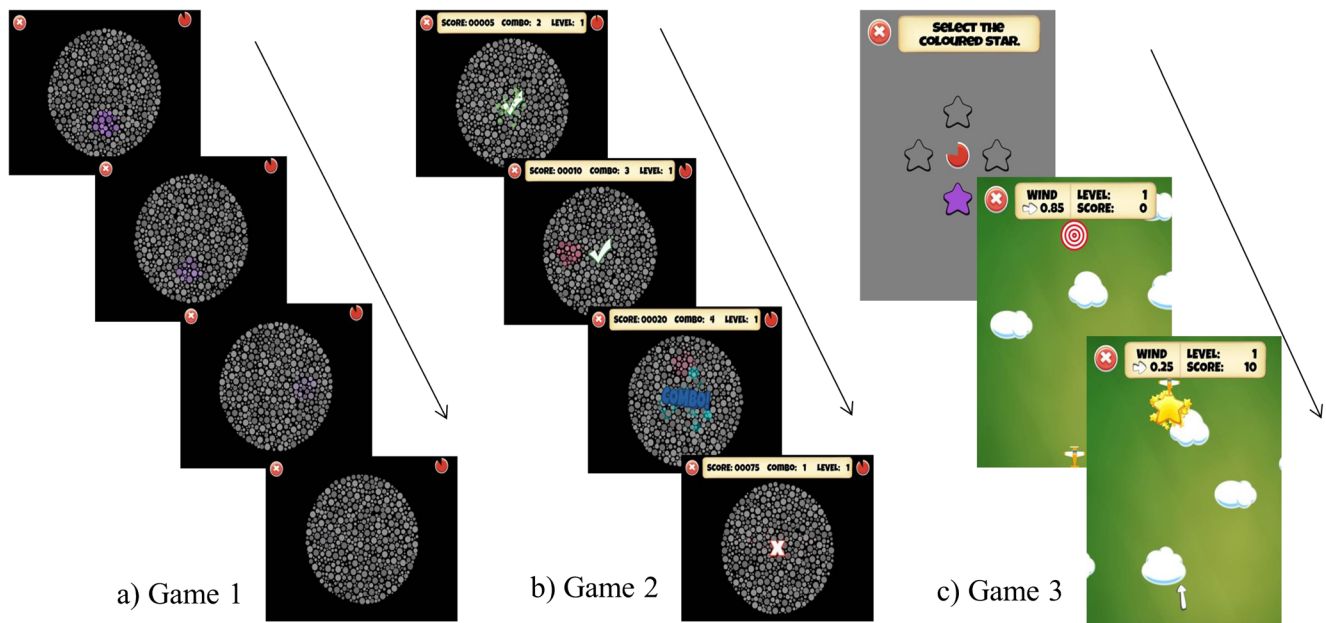
The chromatic contrast thresholds were determined along red–green and blue–yellow color axes (Krauskopf, Williams, & Heeley, 1982) (Fig. 1). The three games were designed for use under normal room illumination (e.g., maximum brightness settings for the iPad were stipulated and in an environment where reflections are minimized and with a battery level of >7%), in recognition of the fact that a portable test will be played under different indoor lighting sources. The testing was performed with the tablet computer held at a viewing



**Fig. 1** The stimuli along red (plus symbols) – green (cross symbols) and blue (triangle symbols) – yellow (circle symbols) cardinal axes (for tablet computer's gamut) shown on CIE 1976 Chromaticity diagram. The overlaid black triangle represents the color gamut of the tablet computer. Letters RGB represent the red, green, and blue ends of the color gamut

distance of 30 cm (participants were instructed to hold and maintain the tablet computer at this viewing distance), where the stimulus (color patch in games 1 and 2 and a colored star in game 3) subtends a visual angle of 3°. The stimulus was presented in a four-alternative locations forced choice manner (4AFC: up, down, right, and left). The participants' task was to detect the location of the colored patch out of the possible four locations amidst the background and indicate its location on the touch-screen display with a tap of the finger within 5s. If a response was not made within 5s, the non-response was regarded as an incorrect response. This relatively long stimulus duration was provided so that inter-individual differences in rapidity of eye-hand coordination would have minimal impact on the results. The games employed a 1-up–2-down staircase procedure (Levitt, 1971) to determine chromatic contrast thresholds. The chromatic contrast of the stimulus was increased with each incorrect/no response (1-up) and decreased after two consecutive correct responses (2-down). A range of 95 to 127 stimuli levels were available for different color directions and the stimulus level of 83 was used as the starting stimulus for all the color directions. The initial stimulus has a relatively high saturation and then the saturation decreases as the test progresses and the subject returns correct answers. From the initial stimulus, the step size was 16 until an incorrect response was recorded. The step size was then halved on every reversal. After every reversal, a control stimulus known to be suprathreshold by virtue of being highly saturated was presented to keep the participant more attentive as it was reasoned that numerous near threshold presentations may be discouraging for the participant and they may need to guess. However, the responses to these control steps were not counted in determining the staircase. A total of eight staircase reversals were recorded and the average of the last four reversals (with the smallest step size of 1) was used to estimate the thresholds (in  $u'v' \times 10^{-4}$  chromaticity coordinates). The number of reversals at the smallest step size was designed to be >1 and <8 which may result in inaccurate thresholds in case of lapses in attention, therefore a compromise value of 4 was selected to determine the thresholds. Figure 2 shows the sequence of game play for three tablet computer-based games. Games 1, 2, and 3 were called *Color detective*, *Color combo rush* and *Flying ace*, in that order. A description of each game is given below.

Game 1, the “*Color detective*” game (Fig. 2a), was designed to be a clinical version of the vision test, without gaming elements. The design of the test was informed by a child-friendly version of the CCT test (Goulart et al., 2008) with the background and stimulus composed of circles of varying size with small variations of added luminance noise ( $\pm 15\%$  of the given RGB units). This luminance noise of  $\pm 15\%$  was considered more than enough to mask any luminance cues from the iPad displays (Bodduluri et al., 2016). The test stimulus was a roughly circular amorphous patch



**Fig. 2** The tablet computer-based application showing the sequence of game play for the three games

which differs in chromaticity from the gray background and thus the luminance artefact may be relatively smaller. Game 2, the “Color combo rush” (Fig. 2b), was a gamified version of the clinical version of the test (game 1). Thus, it was designed to be similar to game 1 in terms of visual stimulus and task, with added fun elements such as feedback (correct/wrong), scores, and sound to facilitate self-administration, engagement, and enjoyment (Fig. 2b).

Game 3, the “Flying ace” (Fig. 2c), was a complete game which included the vision test. The design of the visual stimulus differed in shape and appearance, and the task was different from games 1 and 2. No luminance noise was employed as within the region of placement of the stimuli for the psychophysical task the variations in luminance were insignificant (Bodduluri et al., 2016). Unlike games 1 and 2, game 3 was a two-part game. In the first part, the chromatic contrast thresholds were assessed using an “odd one out” task: Four stars were presented in a diamond configuration, on a gray background, in the center of the tablet computer’s display. Three of the stars were filled with the same background gray but the fourth star was the target stimulus (colored). Thus it was a four-alternative forced choice, but the choice was to pick the “odd one out” (colored star) from the four stars. Thereafter, phase 2 began where the task was to use a flicking action on the touchscreen to launch a plane towards a target. The target was stationary, but wind speed and direction could vary and clouds could obscure the view (Fig. 2c). The task was designed to have numerous variations and not be too demanding on cognitive ability. Points were awarded if the plane intersected the target. In game 3, the color assessment and game aspects were separated in order for the game to look

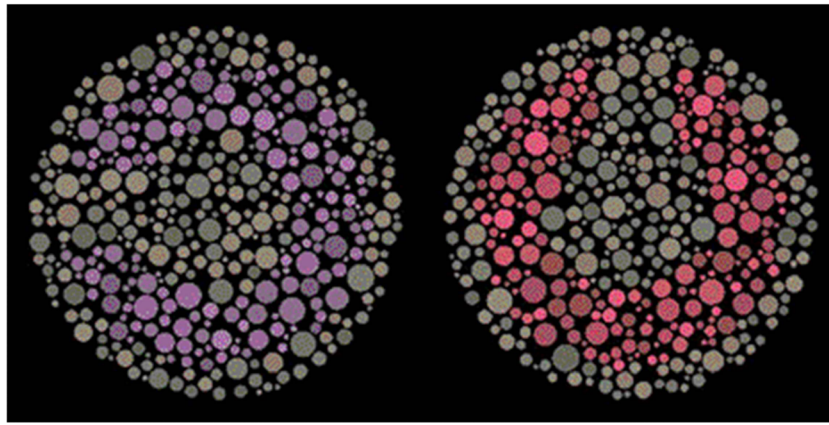
like a game without compromising the requirements of the vision assessment.

#### *Cambridge Colour Test: Trivector test*

The Cambridge Colour Test (CCT) Trivector test, CCT v1.5 (Cambridge Research Systems (CRS) Ltd., Rochester, UK) was run on a cathode ray tube (CRT) monitor (HP p1230, HP, UK) that was calibrated using the manufacturer-specified guidelines provided by CRS (ColorCAL II Colorimeter, VSG 72.12.40F1) prior to the experiment.

The CCT Trivector test stimulus, a Landolt-like “C” (Fig. 3), is formed by randomly distributed gray circles of varying size and luminance. The luminance of the circles was randomly set at any of six luminance levels between 8 and 18  $\text{cd/m}^2$  and the chromatic contrast of the Landolt “C” was varied relative to the fixed gray background ( $u' = 0.198$ ,  $v' = 0.469$ ) (Mollon & Regan, 2000). The color thresholds were determined along three test axes in color space along the protan, deutan, and tritan color confusion lines (Fig. 4). The CCT test was performed in a darkened room with the participant at a viewing distance of 3 m, such that the gap in the “C” subtended a visual angle of  $1^\circ$ . The stimulus was presented in a four-alternative forced choice manner. The participants’ task was to identify the orientation of the gap in the Landolt “C,” and enter the response using a response box (CT6, CRS) within 5s. There were some random control trials (high in color saturation and easily identifiable stimuli) during the testing to check for attention of the participant. This test uses a staircase procedure to obtain the chromatic contrast





**Fig. 3** The target stimulus (Landolt C) of the Cambridge Colour Test (CCT) Trivector test with different chromaticity from background

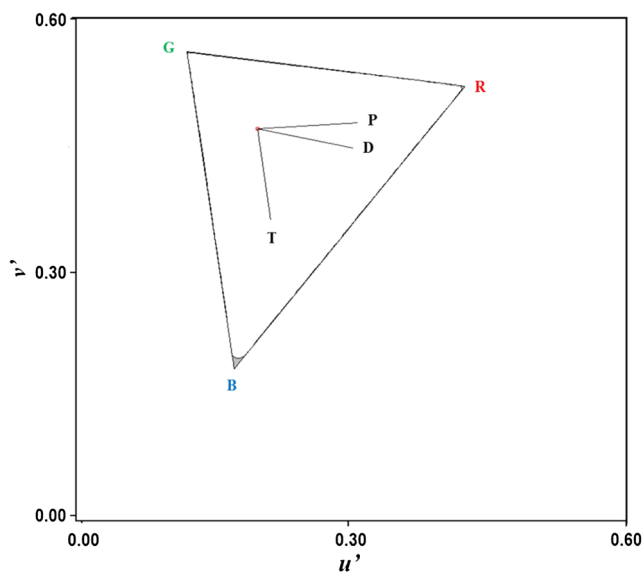
thresholds and the staircases for three test axes were interleaved and run simultaneously in random manner. The staircase began with a stimulus of high saturation and the chromaticity was then varied, i.e. the chromatic contrast of the stimulus was reduced after a correct response and increased following an incorrect or no response. After six staircase reversals, color thresholds were computed as the average of chromaticities corresponding to the reversals (Mollon & Regan, 2000).

Analysis of our pilot data showed no significant differences between monocular and binocular assessment of chromatic contrast thresholds ( $F_{(1,4)} = 1.55$ ,  $p = 0.28$ ) and thus all the testing was performed binocularly with the habitual correction (wherever applicable). The order of testing was randomized (for both the CCT Trivector test and for the three games) to minimize any learning and fatigue effects. A total of 100

participants were recruited and of those, repeatability measurements were obtained on 93 participants (71 participants on the same day with a 20- to 30-min interval between the administrations – intrasession, and the remaining 22 participants had the second set of measurements on a different day – intersession). As there were no significant differences between the intra- and intersession measurements, all the data was combined to analyze the repeatability. The chromatic contrast thresholds were reported as the  $\Delta u'v' \times 10^{-4}$ .

#### Data analysis

As the data was not normally distributed, descriptive statistics were given as the median and interquartile range (IQR) wherever applicable. Tolerance intervals were calculated to report the lower and upper limits of the normative values for the three games as well as for the CCT Trivector test. The protan and deutan chromatic contrast thresholds from the CCT Trivector test were averaged to compare with the red–green thresholds from tablet computer-based games while the tritan thresholds were directly compared with the blue–yellow thresholds. Friedman analysis of variance (ANOVA) with *post-hoc* Wilcoxon signed-rank test was used to compare the chromatic contrast thresholds from the tablet computer-based app with that of the CCT Trivector test. The corresponding Bonferroni-adjusted significance levels were set at  $p < 0.008$  for pairwise comparisons. Wilcoxon signed-rank test was used to test the repeatability of the tests and a complementary measure of agreement was given through the Bland–Altman analysis (Bland & Altman, 1999; Carkeet & Goh, 2016). Bland–Altman plots were plotted using mean ( $\pm 95\%$  limits of agreement (*LoA*): calculated using  $1.96 \times \text{SD}$ ) and median ( $\pm 95\%$  *LoA*: calculated using percentiles) for normally and non-normally distributed differences, respectively. The repeatability of the tablet computer-based app was compared with the CCT Trivector test by determining the 95% *LoA*. Statistical Package for Social Science (SPSS) version 22 and



**Fig. 4** The three test axes (P: Protan, D: Deutan, and T: Tritan) of the Cambridge Colour Test (CCT) Trivector test. Letters RGB represent the red, green, and blue ends of the color gamut

Minitab version 17 were used for statistical analyses and for determining the tolerance interval for tablet computer-based games, respectively.

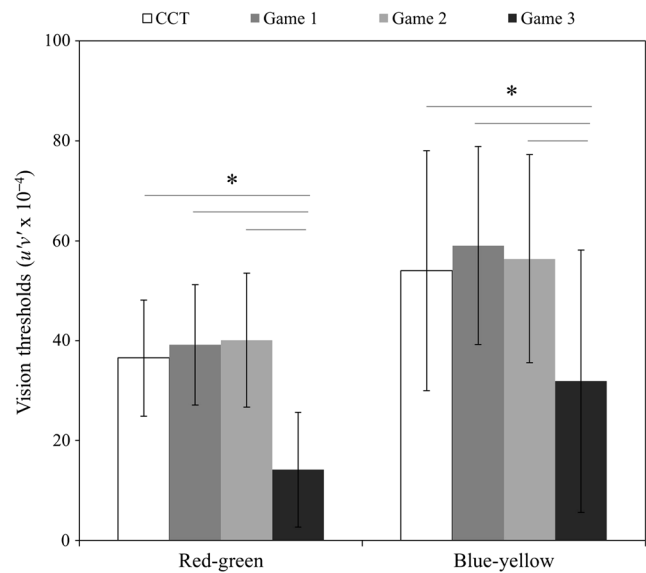
## Results

### Tolerance intervals for tablet computer-based application

Tolerance intervals are defined as the interval within which the given percentage of population falls with a given probability. In the present study, non-parametric tolerance intervals were determined and were given for 95% of the population with a probability of 95%. The chromatic contrast thresholds were measured for red–green and blue–yellow for all three tablet computer-based games and for the CCT Trivector test. The median (IQR) red–green and blue–yellow chromatic contrast thresholds and the corresponding tolerance intervals for each of the games in comparison with the CCT are given in Table 1.

### Comparison between CCT Trivector test and the three games

Friedman ANOVA showed an overall significant difference in measured absolute thresholds for the CCT and the three games ( $\chi^2(7) = 413.6, p < 0.001$ ). *Post-hoc* Wilcoxon signed-rank test, with Bonferroni adjusted *p*-value (0.008), showed no significant difference between the CCT Trivector test and games 1 and 2 (game 1: red–green:  $Z = -1.56, p = 0.12$ , blue–yellow:  $Z = -2.33, p = 0.02$ ; game 2: red–green:  $Z = -2.23, p = 0.03$ , blue–yellow:  $Z = -1.40, p = 0.16$ ). However, game 3, due to its design characteristics, showed significantly lower absolute thresholds compared with the CCT Trivector test as well as with the other two games (red–green:  $Z = -8.41, p < 0.0001$ , blue–yellow:  $Z = -6.26, p < 0.0001$ ). Figure 5 shows the median (IQR) absolute chromatic contrast thresholds (for red–green



**Fig. 5** Median (interquartile range) red–green and blue–yellow thresholds for the Cambridge Colour Test (CCT) Trivector test and three games. \*indicates significant difference between the tests ( $p < 0.008$ , *post-hoc* Wilcoxon signed-rank test)

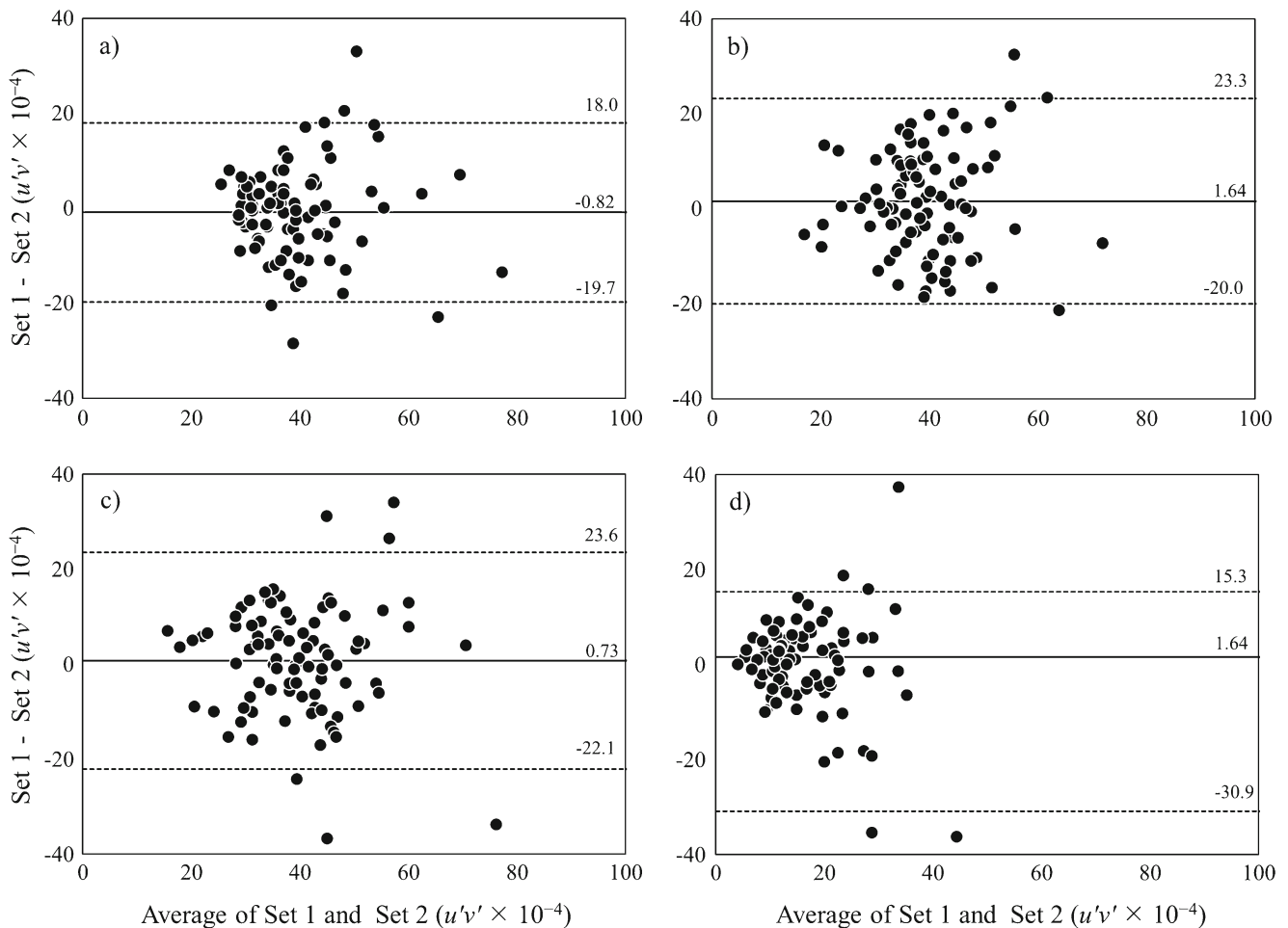
and blue–yellow) for the CCT Trivector test and three tablet computer-based games.

### Repeatability of CCT Trivector test and tablet computer-based application

A total of 93 participants had two sets of measurements obtained using the CCT Trivector test and three games. The Wilcoxon signed-rank test showed no significant difference between two sets of measurements ( $p > 0.05$ ). Though there is no significant difference between two sets of measurements, the 95% *LoA* for game 3 blue–yellow thresholds were higher compared with the other two games and the CCT Trivector test indicating a greater variability (poor repeatability) in the measurements. Figures 6 and 7 show the Bland–Altman plots for repeatability of the CCT Trivector test and 3 tablet computer-based games for red–green and blue–yellow, respectively.

**Table 1** The median (interquartile range) and tolerance intervals for the three games and for the Cambridge Colour Test (CCT) Trivector test. The thresholds were reported as  $u'v' \times 10^{-4}$

	Red–Green	Tolerance interval		Blue–Yellow	Tolerance interval	
		Lower limit	Upper limit		Lower limit	Upper limit
CCT	36.5 (11.6)	24.5	73.0	54.0 (24.0)	26.0	127
Game 1	39.2 (12.1)	16.2	104	59.1 (19.8)	26.5	98.9
Game 2	40.1 (13.4)	14.2	88.8	56.4 (20.9)	22.2	106
Game 3	14.2 (11.5)	4.05	52.4	31.9 (26.3)	2.68	90.6



**Fig. 6** Repeatability of (a) the Cambridge Colour Test (CCT) Trivector test, (b) game 1, (c) game 2, and (d) game 3 for red–green thresholds. The mean difference and the 95% *LoA* for (a), (b), (c) and median difference

and 95% *LoA* for (d) are shown by straight and dashed lines with the values, respectively

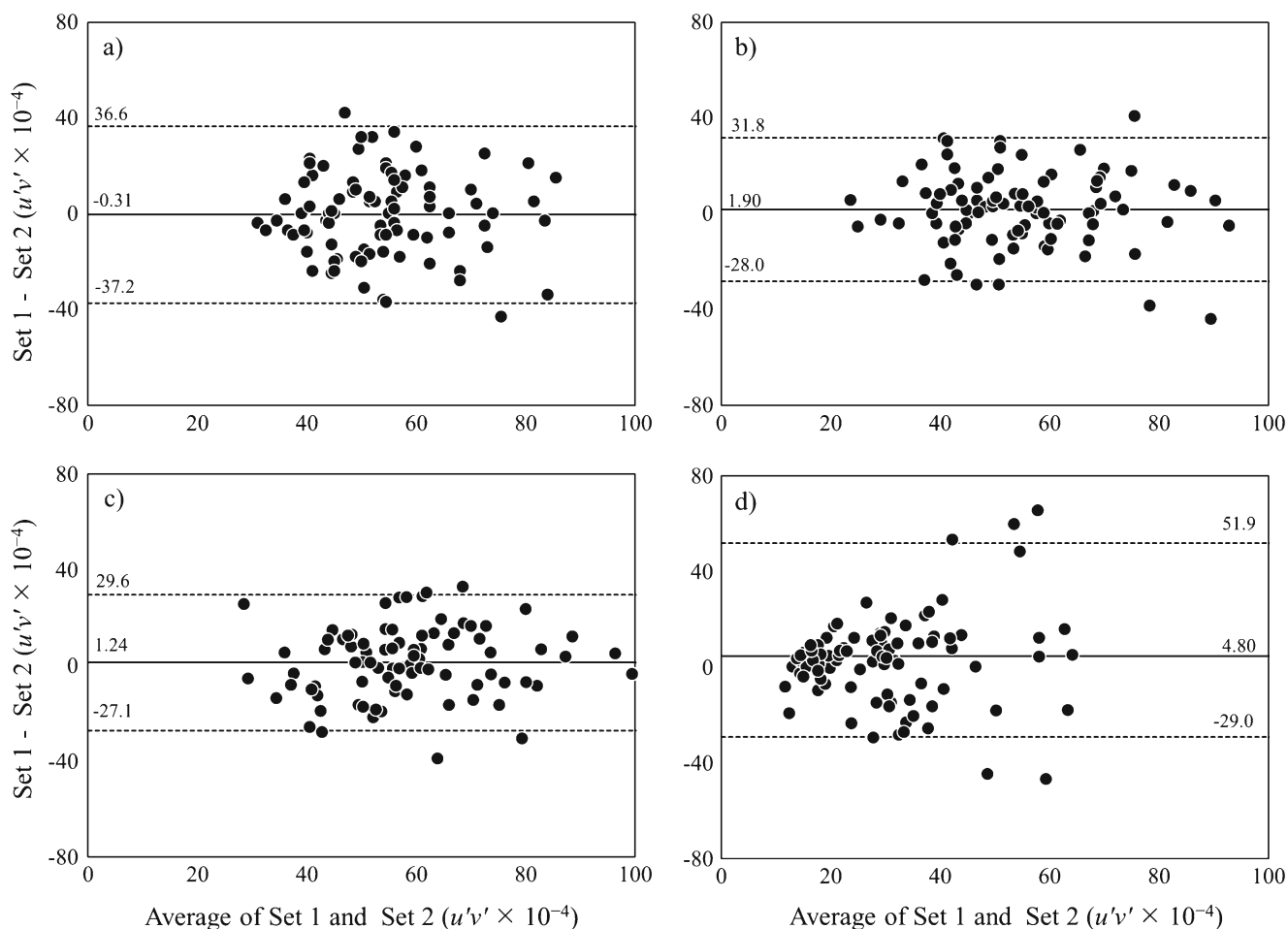
## Discussion

The current study documented adult normative limits (tolerance intervals), including the repeatabilities, for a custom-designed tablet computer-based app, which had three games, in comparison with an established test, the CCT Trivector test. The tolerance intervals for the three games and the CCT are lower than the published norms in the CCT Manual (Mollon & Regan, 2000). However, the tolerance intervals determined in the current study were slightly above the upper limit for the red–green thresholds, for games 1 and 2, than those reported by Ventura, Silveria, et al. (2003) ( $76 u'v' \times 10^{-4}$  in Ventura et al. versus game 1:  $104 u'v' \times 10^{-4}$  and game 2:  $88.8 u'v' \times 10^{-4}$ ).

The chromatic contrast thresholds obtained using games 1 and 2 were comparable with that of the published norms of the CCT Trivector test (Ventura DF, 2003) and also with those obtained using the children friendly version of the CCT Trivector test (Goulart et al., 2008) in adults. This is likely to be due to the similarities in the test design and stimulus

characteristics such as the pseudoisochromatic design, presence of luminance noise, and the similar stimulus presentation, i.e. the colored patch (in the case of the child-friendly version of CCT). It must be noted that although these similarities suggest that games 1 and 2 findings should be more similar to the CCT findings than to game 3, it does not necessarily follow that games 1, 2, and the CCT are expected to behave identically. That is because there are also some differences between these tests, including the stimulus size, test distance, the psychophysical methods (such as staircase characteristics) the color axes used to determine chromatic thresholds, and the stimulus durations that were employed.

There was a significant difference between the chromatic contrast thresholds obtained using game 3 (lower thresholds) and that of the CCT Trivector test and the other two games. This was expected as the test design and the stimulus characteristics of game 3 were different compared with the other games and the CCT test; most notably by the use of a black outline for the stars, the absence of luminance noise to mask artefactual luminance clues, and the “odd one out” task



**Fig. 7** Repeatability of (a) the Cambridge Colour Test (CCT) Trivector test, (b) game 1, (c) game 2, and (d) game 3 for blue–yellow chromatic contrast thresholds. The mean difference and the 95% LoA for (a), (b), (c)

and median difference and 95% LoA for (d) are shown by straight and dashed lines with the values, respectively

procedure. Moreover, luminance noise can affect the appearance of a stimulus and in turn the estimated visual thresholds. Therefore, the lower thresholds obtained using game 3 could possibly be due to the absence of luminance noise.

In order to understand the variability in the spread of chromatic contrast thresholds on repeated administrations, the 95% LoA for all the tests were compared. The 95% LoA of the three games and the CCT Trivector test were comparable, except for the blue–yellow thresholds using game 3 which showed greater variability indicating poor repeatability. The test-retest variability is relatively high compared to the spread of scores; this may be a drawback. However, as this finding is similar to the CCT Trivector test (J.D. Mollon, 2000), with the exception of the game 3 blue–yellow test, and given that the CCT Trivector test has been used to report statistically significant differences in color vision in people with diabetes (Gualtieri, Feitosa-Santana, Lago, Nishi, & Ventura, 2013), above community levels of exposure to occupational solvents (Costa et al., 2012) and in smokers (Fernandes & Santos,

2017), it may be reasoned that the tablet-computer-based games have the same potential to provide clinically significant measures.

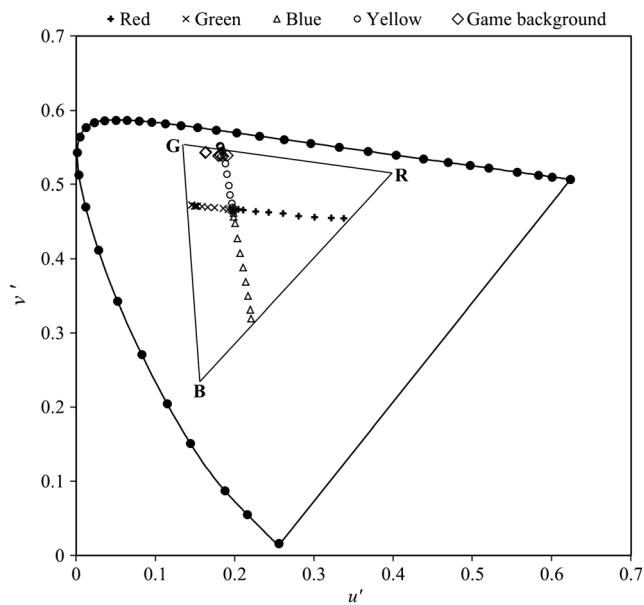
Game 3 was also observed to result in blue–yellow thresholds that increase in difference between two administrations with increasing thresholds, i.e. the data points in the Bland–Altman plots (Figs. 6d and 7d) appear to diverge with higher thresholds for game 3. The possible reasons for this wide-spread 95% LoA and the greater variability on repeated testing in game 3 (especially for blue–yellow) include the design characteristics of the test (including the incorporation of a plane-flying component) or a possible learning effect (explained later) with repeated testing. To clarify the situation, pilot data from a small sample of participants who had played game 3 both with and without the plane-flying component (game part of the third game) was reviewed and analyzed. The analysis showed no significant difference between repeated testing for the red–green thresholds, but there was a significant difference in the thresholds for blue–yellow, i.e. the



thresholds were lower when there was no game part. On further evaluation of the design of the game part, it was observed that the color of the background in the game part (please see the color of the second and third panels in Fig. 2c) was similar to the colors used in the test stimuli and was clustered around the yellow direction used in the actual color vision test (Fig. 8). Therefore, it is likely that the color of the game background adapted the blue–yellow system of users variably, contributing to greater variability in thresholds. This finding highlights the importance of ensuring that the gaming components cannot interfere with the psychophysical task. This finding also indicates that while the cognitive load is higher for game 3 than games 1 and 2, this did not appear to affect the participants as the red–green thresholds did not show a difference whether the game part was included or excluded.

The red–green thresholds finding also suggests that it is possible to have a psychophysical task nested within where the participant has to mentally switch between two tasks (the game and the psychophysical task) regularly and demanding cognitive tasks in the game component (in this case, participants were required to intersect the target with the plane and make judgments about speed, trajectory, or anticipated disruption due to wind speed within the game environment) without unduly affecting psychophysical thresholds.

Although there was no significant difference between the first and second sets of measurements, hence the results were repeatable, the chromatic contrast thresholds were lower on



**Fig. 8** The stimuli along red (plus symbols) – green (cross symbols) and blue (triangle symbols) – yellow (circle symbols) cardinal axes (for tablet computer’s gamut) shown on CIE 1976 Chromaticity diagram. The “diamond” symbols overlapping the circles in the yellow direction represent the chromaticities of the background colors used in the plane flying component of game 3. The overlaid black triangle represents the color gamut of the tablet computer. Letters RGB represent the red, green, and blue ends of the color gamut

the second administration of the test for all types of thresholds tested. This may represent a small learning effect. This improvement ranged from 2% to 10% with a mean global improvement of 4% (corresponds to  $1.5 u'v' \times 10^{-4}$ ) in the second administration. However, this difference is not clinically significant and is well within the 95% *LoA* of the repeatability.

We found from the analysis that small color differences such as chromatic contrast thresholds can be measured using the iPad mini retina display device along color axes and the results were comparable to those tested using the CCT. The increasing sophistication of tablet devices and their display resolutions has enabled the development of apps that can be used in the assessment of visual function (Dorr et al., 2013; Kollbaum et al., 2014; Rodriguez-Vallejo et al., 2015). Considering the future, it is likely that the development of these kinds of vision apps will further enable testing with greater technical sophistication such as within a gaming environment. This study’s findings also highlight that it is important to consider how more sophisticated forms of game tasks and psychophysical tasks may interact, particularly how measures of psychophysical thresholds may be affected.

Although this study provides a portable tablet computer-based vision testing app, it has certain limitations as discussed in here. It may not be possible to use this app interchangeably with other generations of tablet computers from the same or other manufacturers, or with other tablet computer-operating systems (e.g. Android or Windows). This is due to different manufacture specifications and variations in display screen technologies of other tablet computers and their color gamut sizes. For example, as reported by Dain, Kwan, and Wong (2016), a single stimuli look-up table cannot be used for different models of smart phones from the same manufacturer (for iPhone 4 s and iPhone 5) as this may lead to significant reproduction errors due to the various factors as listed above. Therefore, it is advisable that care must be taken to separately calibrate and develop device-specific stimuli look-up tables when applying this vision test for use on other devices.

In summary, the three tablet computer-based games, with the exception of the blue–yellow component of game 3, have been found to provide estimates of chromatic contrast thresholds, in a self-administrable and portable format. Additionally, games 1 and 2 yield visual thresholds comparable with the CCT Trivector test where the test and background stimulus are of similar shape and contain luminance noise (games 1 and 2). Their portable and self-administrable design will allow these tablet computer-based games to be used to assess chromatic contrast thresholds outside the research laboratory or in a routine clinical setting. The games presented in the current study were designed to assess normal age-related variations and acquired deficits in chromatic contrast sensitivity, not to detect and diagnose congenital color vision deficiencies. Further work would still be required to understand how tablet

computer displays may be used to detect and diagnose congenital color vision deficiencies, which would require the stimuli to be presented along color confusion axes similar to that in the CCT (J.D Mollon, 2000).

### Compliance with ethical standards

**Competing interests** None of the authors has any potential competing interests.

### References

- Abramov, I., Hainline, L., Turkel, J., Lemerise, E., Smith, H., Gordon, J., & Petry, S. (1984). Rocket-ship psychophysics. assessing visual functioning in young children. *Investigative Ophthalmology & Visual Science*, 25(11), 1307–1315.
- Anderson, K., Burford, O., & Emmerton, L. (2016). Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences. *PLoS ONE*, 11(5), e0156164. doi:10.1371/journal.pone.0156164
- Aslam, T. M., Murray, I. J., Lai, M. Y. T., Linton, E., Tahir, H. J., & Parry, N. R. A. (2013). An assessment of a modern touch-screen tablet computer with reference to core physical characteristics necessary for clinical vision testing. *Journal of the Royal Society Interface*, 10(84), 20130239. doi:10.1098/rsif.2013.0239
- Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8(2), 135–160.
- Bodduluri, L., Boon, M. Y., & Dain, S. J. (2016). Evaluation of tablet computers for visual function assessment. *Behavior Research Methods*, 1–11. doi:10.3758/s13428-016-0725-1
- Carkeet, A., & Goh, Y. T. (2016). Confidence and coverage for Bland–Altman limits of agreement and their approximate confidence intervals. *Stat Methods Med Res*, 0962280216665419. doi:10.1177/0962280216665419
- Costa, T. L., Barboni, M. T. S., de Araujo Moura, A. L., Bonci, D. M. O., Gualtieri, M., de Lima Silveira, L. C., & Ventura, D. F. (2012). Long-term occupational exposure to organic solvents affects color vision, contrast sensitivity and visual fields. *PLoS ONE*, 7(8), e42961.
- Cranwell, M. B., Pearce, B., Loveridge, C., & Hurlbert, A. C. (2015). Performance on the Farnsworth-Munsell 100-Hue Test Is Significantly Related to Nonverbal IQFM100 Test Significantly Related to NVIQ. *Investigative Ophthalmology & Visual Science*, 56(5), 3171–3178. doi:10.1167/iovs.14-16094
- Dain, S. J. (2004). Clinical colour vision tests. *Clinical & Experimental Optometry*, 87(4-5), 276–293. doi:10.1111/j.1444-0938.2004.tb05057.x
- Dain, S. J., Kwan, B., & Wong, L. (2016). Consistency of color representation in smart phones. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 33(3), 300–305. doi:10.1364/josaa.33.00a300
- Dain, S. J., & Ling, B. Y. (2009). Cognitive abilities of children on a gray seriation test. *Optometry and Vision Science*, 86(6), E701–E707. doi:10.1097/OPX.0b013e3181a59d46
- Dorr, M., Lesmes, L. A., Lu, Z.-L., & Bex, P. J. (2013). Rapid and reliable assessment of the contrast sensitivity function on an iPad. *Investigative Ophthalmology & Visual Science*, 54(12), 7266–7273. doi:10.1167/iovs.13-11743
- Feitosa-Santana, C., Oiwa, N. N., Paramei, G. V., Bimler, D., Costa, M. F., Lago, M., ... Ventura, D. F. (2006). Color space distortions in patients with type 2 diabetes mellitus. *Visual Neuroscience*, 23(3–4), 663–668. doi:10.1017/s0952523806233546
- Fernandes, T. M. d. P., & Santos, N. A. d. (2017). Comparison of color discrimination in chronic heavy smokers and healthy subjects [version 1; referees: Awaiting peer review]. *F1000Research*, 6(85). doi:10.12688/f1000research.10714.1
- Goulart, P. R., Bandeira, M. L., Tsubota, D., Oiwa, N. N., Costa, M. F., & Ventura, D. F. (2008). A computer-controlled color vision test for children based on the Cambridge Colour Test. *Visual Neuroscience*, 25(3), 445–450. doi:10.1017/s0952523808080589
- Greenstein, V. C., Hood, D. C., Ritch, R., Steinberger, D., & Carr, R. E. (1989). S (blue) cone pathway vulnerability in retinitis pigmentosa, diabetes and glaucoma. *Investigative Ophthalmology & Visual Science*, 30(8), 1732–1737.
- Gualtieri, M., Feitosa-Santana, C., Lago, M., Nishi, M., & Ventura, D. F. (2013). Early visual changes in diabetic patients with no retinopathy measured by color discrimination and electroretinography. *Psychology & Neuroscience*, 6, 227–234.
- Ishihara, S. (1917). *Test for colour-blindness*. Tokyo: Hongo Harukicho.
- Jacobs, G. H. (1993). The distribution and nature of colour vision among the mammals. *Biological Reviews*, 68(3), 413–471. doi:10.1111/j.1469-185X.1993.tb00738.x
- Kollbaum, P. S., Jansen, M. E., Kollbaum, E. J., & Bullimore, M. A. (2014). Validation of an iPad test of letter contrast sensitivity. *Optometry and Vision Science*, 91(3), 291–296. doi:10.1097/oxp.000000000000158
- Krauskopf, J., Williams, D. R., & Heeley, D. W. (1982). Cardinal directions of color space. *Vision Research*, 22(9), 1123–1131. doi:10.1016/0042-6989(82)90077-3
- Kurtenbach, A., Schiefer, U., Neu, A., & Zrenner, E. (1999). Development of brightness matching and colour vision deficits in juvenile diabetics. *Vision Research*, 39(6), 1221–1229. doi:10.1016/s0042-6989(98)00214-4
- Lakowski, R. (1969). Theory and practice of colour vision testing: A review. Part 2. *British Journal of Industrial Medicine*, 26(4), 265–288. doi:10.1136/oem.26.4.265
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2), 467–477.
- MailOnline-Australia. (2015). Can YOU spot the odd one out? KukuKube puts colour vision to the test. Retrieved from <http://www.dailymail.co.uk/sciencetech/article-3033455/How-good-colour-vision-KukuKube-app-tests-ability-subtle-differences-shade-leave-cross-eyed.html>
- Mäntyjärvi, M. (1987). An Evaluation of the Standard Pseudoisochromatic Plates (SPP 1) in Clinical Use. In G. Verriest (Ed.), *Colour Vision Deficiencies VIII* (pp. 125–131). Dordrecht: Springer Netherlands.
- Melamad, A., Hagstrom, S., & Traboulsi, E. (2004). Color vision testing. *Ophthalmic Genetics*, 25(3), 159–187. doi:10.1080/13816810490498341
- Mollon, J. D., & Regan, B. C. (2000). *Cambridge Colour Test Handbook*. (Cambridge Research Systems Ltd., 2000), Version 1.1.
- Mulligan, J. B. (2013). Rapid assessment of contrast sensitivity with mobile touch-screens. *Journal of Vision*, 13(9), 270. doi:10.1167/13.9.270
- Nguyen, L. C., Do, E. Y.-L., Chia, A., Wang, Y., & Duh, H. B.-L. (2014). *DoDo game, a color vision deficiency screening test for young children*. Paper presented at the Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, Toronto, Ontario, Canada.
- O'Neill-Biba, M., Sivaprasad, S., Rodriguez-Carmona, M., Wolf, J. E., & Barbur, J. L. (2010). Loss of chromatic sensitivity in AMD and diabetes: A comparative study. *Ophthalmic and Physiological Optics*, 30(5), 705–716. doi:10.1111/j.1475-1313.2010.00775.x
- Rodriguez-Vallejo, M., Remon, L., Monsoriu, J. A., & Furlan, W. D. (2015). Designing a new test for contrast sensitivity function

- measurement with iPad. *Journal of Optometry*, 8(2), 101–108. doi: [10.1016/j.optom.2014.06.003](https://doi.org/10.1016/j.optom.2014.06.003)
- Seshadri, J., Christensen, J., Lakshminarayanan, V., & Bassi, C. J. (2005). Evaluation of the new web-based “Colour Assessment and Diagnosis” test. *Optometry and Vision Science*, 82(10), 882–885. doi: [10.1097/01.opx.0000182211.48498.4e](https://doi.org/10.1097/01.opx.0000182211.48498.4e)
- Turpin, A., Lawson, D. J., & McKendrick, A. M. (2014). PsyPad: A platform for visual psychophysics on the iPad. *Journal of Vision*, 14(3), 16. doi: [10.1167/14.3.16](https://doi.org/10.1167/14.3.16). 11–17.
- Ventura, D., Costa, M., Gualtieri, M., Nishi, M., Bernick, M., Bonci, D., & De Souza, J. (2003). Early vision loss in diabetic patients assessed by the Cambridge Colour Test. *Normal and defective colour vision*, 395–403. doi: [10.1093/acprof:oso/9780198525301.003.0042](https://doi.org/10.1093/acprof:oso/9780198525301.003.0042)
- Ventura, D., Silveira, L., Rodrigues, A., De Souza, J., Gualtieri, M., Bonci, D., & Costa, M. (2003). Preliminary norms for the Cambridge Colour Test. In J. D. Mollon, J. Pokorny, & K. Knoblauch (Eds.), *Normal and defective colour vision* (pp. 331–339). Oxford: Oxford University Press.