



City Research Online

City St George's, University of London

Citation: Evans, B. E. W., Rodriguez-Carmona, M., Llapashtica, E. & Barbur, J. L. (2025). Color screening outcomes on uncalibrated “home” displays—assessment of variability in sRGB color standard parameters on commercial visual displays. *Journal of the Optical Society of America A*, 42(5), B289-B295. doi: 10.1364/josaa.544986

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/34848/>

Link to published version: <https://doi.org/10.1364/josaa.544986>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Color screening outcomes on uncalibrated 'home' displays – assessment of variability in sRGB color standard parameters on commercial visual displays

Benjamin E W Evans [1], ORCID: 0000-0001-6611-4500
Marisa Rodriguez-Carmona [1], ORCID: 0000-0002-7366-6159
Emsal Llapashtica [1], ORCID: 0000-0003-1941-9971
John L Barbur [1] ORCID: 0000-0002-2187-5004

[1] The Henry Wellcome Laboratories for Vision Science, Centre for Applied Vision Research, School of Health and Medical Sciences, City St George's, University of London, UK.

Abstract

The Colour Vision Screener (CVS) test was evaluated for use on visual displays that support the Red Green Blue (sRGB) color space standard. When used on commercial displays in sRGB color mode, the CVS had a sensitivity and specificity (95 percent confidence intervals) of 0.55 (0.46 – 0.64) and 1.00 (0.98 – 1.00). With calibrated displays all color deficient observers failed the CVS, but many of these, some with severe loss, were able to pass the test on 10 separate sRGB displays. The CVS test can only achieve close to 100% sensitivity and specificity on fully calibrated displays.

Keywords: sRGB, color vision, screening, home testing

25 Introduction

26 At-home vision screening has become an increasingly valuable tool in personalized
27 healthcare, allowing for the early detection of disease and the monitoring of progression. [1–
28 3] Such screening can also play a key role in clinical trials, serving as both an endpoint and
29 interim measurement to enhance patient engagement, reduce drop-out rates, and streamline
30 data collection. [4,5] Many methods of screening for visual acuity and amblyopia [6], and
31 color vision [7–12] (and in some cases all three [13]) have been reported, with screeners
32 designed for at-home use typically designed for self-administration, being completed on a
33 personal computer (either via download or webpage) or using a tablet or smartphone.

34 Screening for normal trichromatic color vision is of value in the detection and monitoring of
35 diseases that affect the visual system, [14–16] and for determining occupational
36 suitability. [17–19] Any anomalous aspect in the series of parallel systems which give rise to
37 human color vision can result in abnormal red/green (RG) and yellow/blue (YB) color
38 vision. [20,21] As a result, in large clinical trials color vision testing serves as a valuable
39 endpoint for evaluating the efficacy of treatments for conditions such as diabetic retinopathy
40 and glaucoma. [16,22,23]

41 Technological advancements and the increased production and availability/accessibility of
42 portable computers present additional opportunities for at-home color vision screening,
43 particularly in populations with limited access to regular eye care. To ensure screening
44 provides accurate and actionable data when carried out remotely one must account for
45 factors including device variability, user adherence to instructions, and display and
46 environmental lighting conditions. Color vision testing presents its own specific demands.
47 Whether reproduced across a range of visual displays or on printed media to measure
48 chromatic sensitivity one must ensure stimuli are produced with the correct chromaticity and
49 luminance to isolate color signals. [24] With visual displays one approach to minimize such
50 differences is to use an established and widely adopted color space, such as the Red Green
51 Blue (sRGB) color space standard defined by the International Electrotechnical
52 Commission. [25]

53 Selecting a standard color space enables adequate color reproduction for normal trichromats
54 and can minimize variability, but in practice variability is not eliminated entirely. Locked
55 brightness and color temperature settings can result in sRGB visual displays failing to
56 accurately replicate the intended colors. In addition, these avoidable factors, inter-monitor
57 variation in gamma values, the spectral power distribution of the three primaries, and
58 deliberate selection of higher or lower maximum display luminances by display manufactures
59 can result in increased variability. It is of value, when in pursuit of at-home color vision

testing, to establish experimentally the variability one can expect within sRGB visual displays and when screening for color vision deficiency on such displays.

Two separate studies were carried out to investigate the newly developed color vision screener (CVS) test. [26] This paper is the second of two manuscripts that describe the results from the two studies. The first manuscript evaluated the test when implemented on fully calibrated visual displays and found that the CVS achieves a high sensitivity and specificity (1.00 and 0.99 respectively) in 430 participants assessed across 9 international centers. The purpose of the second study was to investigate the feasibility of 'at home' color vision screening using the CVS when the display calibration is assumed to correspond to the sRGB color standard. This approach replicates the conditions one would expect if the test were downloaded and completed on personal computers outside of the clinic or laboratory. Secondary aims included an investigation into the effect of changes in background chromaticity upon chromatic sensitivity.

Methods

Twenty visual displays that were representative of typical desktop computer monitors commonly used in home and office settings were procured. These visual displays were either commercially available or actively in use in real-world office settings at the outset of the study. Displays were selected based on their affordability and prevalence in workplace environments, particularly those likely to be issued to employees working from home. Half of the displays were acquired from City St Georges, University of London and half were purchased from online retailers. The displays used either cold cathode backlighting, LED, or WLED backlighting and were set to the sRGB color mode using each monitor's on-screen display (OSD) interface with minimal changes to any other setting to mirror, as closely as possible, the procedure individuals would follow if carrying out the tests at home. The spectral radiance of each primary colour and the display sRGB 'white' were measured over the spectral range 380 to 780 nm using a Konica CS-2000A telespectroradiometer (Konica Minolta Inc). The instrument, defined by the manufacturer's specifications, has a specified wavelength resolution of just under 1 nm with a spectral half bandwidth of 5 nm, a precision of ± 0.3 nm, and a chromaticity accuracy of $x: \pm 0.015$, $y: \pm 0.001$ (stated for standard source A). The spectrum was sampled every nm. Gamma functions for each primary were measured for a subset of monitors due to resource constraints during data collection.

From the set of 20 displays, the 10 visual displays with the most recent manufacturer date were selected to ensure the study best reflected current technological standards (displays 11-20 – all of which employ LED or WLED backlighting) and thirty-two participants were recruited using advertisements posted at City St George's, University of London.

Participants completed the CAD and CVS test on a 10 bit Eizo CS2420 monitor, calibrated using a Konica CS-2000A telespectroradiometer (Konica Minolta Inc), and the downloadable CVS test on the ten sRGB visual displays. The CAD and CVS tests have been described previously [26,27] and the initial CAD and CVS test were always completed prior to testing with the sRGB visual displays. The order of testing for the downloadable CVS on the sRGB visual displays was randomized. Participants who were identified as having color vision deficiency by the CAD test completed additional testing using a Type 1 Nagel anomaloscope. The sensitivity and specificity were calculated for all monitors and 95% confidence intervals (CI) were calculated using the Wilson method. Local research and ethical approval was obtained prior to any participant recruitment, informed consent was obtained from all participants, and the project followed the ethical principles of the Declaration of Helsinki.

Following data collection, an exploratory study was carried out to evaluate the potential impact of measuring chromatic sensitivity at different background chromaticities. Using the CAD test paradigm, thresholds were measured for a healthy 30-year-old participant in 32 orientations (14 RG, 18 YB) for 6 background chromaticities. The background chromaticities included: the four background chromaticities with the largest chromatic deviations from the sRGB standard measured in the 10 monitors used in the main study, the background chromaticity used in the CAD test, and the sRGB background chromaticity, as defined in the sRGB standard. Ellipses were fitted to these thresholds using least square fitting described by Fitzgibbon et al., Halir & Flusser, and Zsombor-Murray. [28–30] This algorithm was implemented in Python 3.11.5. [31] The 95% confidence intervals for the ellipse parameters were calculated and used, in combination with the CAD test thresholds, to determine whether the chromatic thresholds differed significantly from one another when measured at each different background chromaticity.

Results

The radiance of each visual display's primary when measured at the start of the study is shown in Fig. 1, along with the maximum luminance. The luminance of the maximum white ranged from 41 to 282 cd/m² for monitors 1-20 and from 82 to 282 cd/m² for monitors 11-20. No significant differences were found between the primary spectra of all visual displays when measured before and after the study.

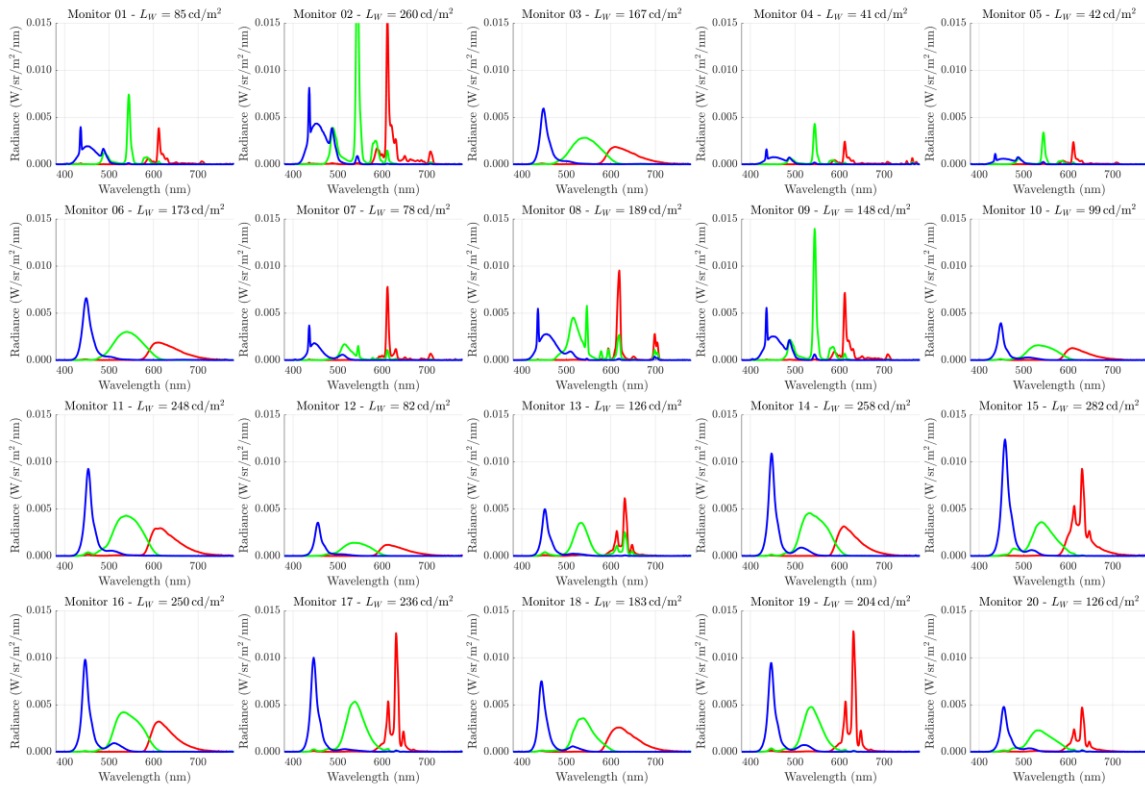


Fig. 1. Initial spectral scans for all 20 monitors set to the sRGB (or standard) color mode measured using a Konica138 CS-2000A telespectroradiometer. The measured luminance of each display is shown in the title for each subplot and the ordinate is fixed to enable a comparison of luminance across all displays.

Table 1. A) The sRGB standard luminance and CIE x,y chromaticity coordinates. B) The photopic luminance of 20 monitors used in the first phase of the study. All monitors were set to their 'standard' or sRGB setting. The luminance and chromaticity of the display 'white' point were measured using a Konica138 CS-2000A telespectroradiometer under the specified sRGB viewing conditions, allowing for comparison to the sRGB standard. Gamma values were only measured for a selection of displays due to resource constraints. The average ΔW in CIE x,y between measurements at the start and end of the study was 0.002 ± 0.001 .

A)

Reference Measure	L _{white} (cd/m ²)	W(x,y)	Gamma
sRGB standard	80	(0.3127,0.3290)	2.2

B)

Monitor No.	L _{photopic} (cd/m ²)	W(x,y)	Gamma	ΔL w.r.t. sRGB (cd/m ²)	White Point chromatic displacement w.r.t sRGB (ΔW CIE x,y)
01	85	(0.287,0.300)	2.1	5	0.039
02	260	(0.321,0.335)	-	180	0.010

03	167	(0.314,0.335)	-	87	0.006
04	41	(0.322,0.344)	-	-39	0.018
05	42	(0.338,0.346)	-	-38	0.031
06	173	(0.308,0.327)	-	93	0.005
07	78	(0.302,0.332)	-	-2	0.011
08	189	(0.339,0.342)	-	109	0.029
09	148	(0.308,0.337)	2.5	68	0.009
10	99	(0.314,0.334)	2.2	19	0.005
11	248	(0.306,0.329)	-	168	0.007
12	82	(0.311,0.329)	2.2	2	0.002
13	126	(0.313,0.333)	2.2	46	0.004
14	258	(0.305,0.332)	-	178	0.008
15	282	(0.312,0.320)	-	202	0.009
16	250	(0.314,0.338)	-	170	0.009
17	236	(0.313,0.330)	-	156	0.001
18	183	(0.315,0.322)	-	103	0.007
19	204	(0.305,0.314)	1.9	124	0.017
20	126	(0.309,0.331)	-	46	0.004

138

139

140

141

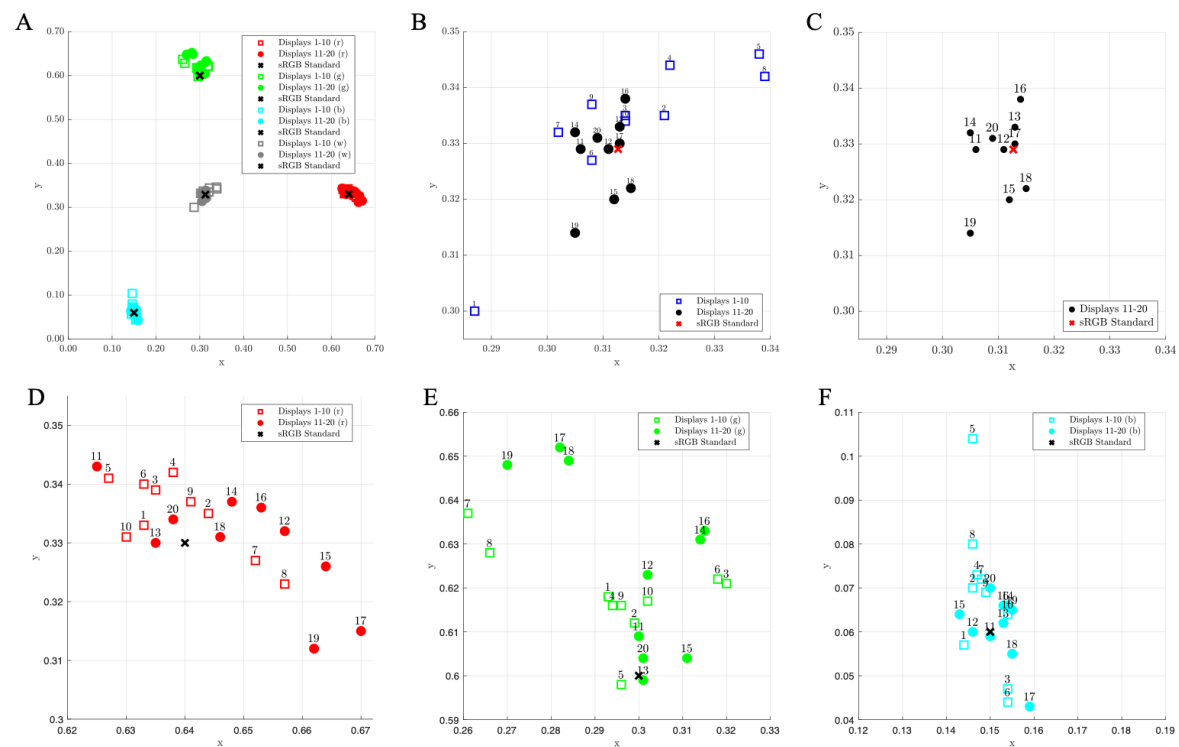
142

143

144

145

Fig. 2 shows the measured CIE x,y chromaticities for each primary color and each monitor and the sRGB standard. Each visual display's white point chromaticity (in CIE x,y coordinates) and the relative difference when compared to the sRGB standard are shown in Table 1. The average displacement in CIE x,y white point chromaticity, across all monitors, in initial and final study measurements was 0.002 ± 0.001 . Gamma values, measured for a subset of the visual displays ranged, across all primaries, from 1.8 to 2.3 with a median [and interquartile range (IQR)] of 2.2 [2.1-2.3].



146

147

148

149

150

Fig. 2. A) The CIE x,y chromaticity coordinates for each of the primaries and white of the 20 visual displays used in the study. B) The CIE x,y chromaticity for the white points of the 20 visual displays and C) the 10 visual displays used for the CVS measurements in the latter part of the study. Plots D), E), and F) show the CIE x,y chromaticity for the red, green, and blue primaries, respectively.

151

152

153

154

155

156

157

158

Participants recruited and assessed on the CVS test included 21 normal trichromats, 8 deutan, and 3 protan participants, as diagnosed by the CAD test. All cases of congenital color deficiency were confirmed using the Nagel anomaloscope (5 deuteranomalous trichromats, 3 deuteranopes, and 3 protanopes), with 100% agreement between the CAD classification and the anomaloscope. Participants ranged from 19 to 69 years of age, with a median age [IQR] of 29 [26-41] years. The median ages for normal trichromats and color deficient were 29 and 30 years respectively. The normal trichromat population included 12 male and 9 female participants, all 11 color deficient were male.

159 All participants were correctly classified by the CVS when carried out using a calibrated
160 visual display. All participants produced correct responses on all displays when presented
161 with 'check' trials (i.e., stimuli defined by a combination of color and luminance contrast
162 signals). As shown in Table 2, all normal trichromats assessed passed the CVS and were
163 correctly classified as having normal RG and YB color vision on all 10 uncalibrated visual
164 displays set up to use the standard or sRGB color mode. Across all visual displays, between
165 25-75% of deuterans and 33-67% of protans passed the CVS test. Three CVDs passed the
166 CVS test on all 10 visual displays and two CVDs failed the CVS test on all 10 visual displays.
167 Treating all 10 uncalibrated visual displays as one group, this corresponds to a RG
168 sensitivity and specificity (and 95% confidence intervals [CI]) of 0.56 (0.46 – 0.64) and 1.00
169 (0.98 – 1.00), respectively for the CVS test's use on uncalibrated sRGB visual displays. The
170 sensitivity of the CVS ranged from 0.27 (0.01 - 0.57) to 0.73 (0.43 - 0.90) on each visual
171 display. For YB stimuli, the CVS specificity was 1.00 (0.99 - 1.00) and the sensitivity could
172 not be determined as no participants with YB loss were assessed.

Table 2. The outcome of 320 CVS tests for 32 participants across 10 uncalibrated visual displays configured with their inbuilt sRGB color mode. A) The proportion of participants with normal color vision, deutan, and protan deficiency who pass the RG component of the CVS on each sRGB display, along with the sensitivity and specificity for each visual display. B) The proportion of RG stimuli correctly identified by the 11 participants with color vision deficiency, sorted by type of color vision deficiency and red/green (RG) CAD threshold. The type of CVD (DA – deuteranomalous, D – deuteranopic, P – protanopic) in color deficiency participants was determined by the CAD test and the Nagel anomaloscope, with 100% agreement between both measures. Cells marked in red indicate cases where participants failed the CVS.

A) The proportion of normal trichromats, deutans, and protans who pass the CVS test on sRGB visual displays

Type of CVD	N	Visual Display									
		11	12	13	14	15	16	17	18	19	20
Normal	21	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Deutan	8	25%	50%	38%	25%	25%	38%	75%	63%	75%	25%
Protan	3	67%	33%	67%	33%	33%	33%	33%	33%	67%	67%
Sensitivity		0.64	0.55	0.55	0.73	0.73	0.64	0.36	0.45	0.27	0.64
Specificity		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

B) The proportion of RG CVS stimuli correctly identified by CVD participants

CVD	Age (years)	CAD Thresholds		Visual Display									
		RG	YB	11	12	13	14	15	16	17	18	19	20
DA	21	3.42	0.82	70%	73%	68%	66%	73%	70%	75%	75%	91%	73%
DA	40	6.17	0.92	52%	64%	70%	50%	66%	61%	82%	59%	59%	59%
DA	23	9.65	0.96	82%	82%	84%	75%	89%	82%	84%	89%	84%	84%
D	26	15.85	0.74	91%	95%	98%	98%	100%	95%	100%	98%	95%	98%
D	34	15.95	0.93	66%	89%	75%	70%	73%	80%	77%	84%	86%	86%
DA	22	16.40	1.04	61%	64%	61%	70%	66%	61%	57%	77%	82%	70%
DA	46	23.96	1.00	70%	77%	68%	68%	68%	64%	75%	73%	80%	70%
D	50	26.47	0.98	50%	64%	57%	61%	45%	66%	64%	64%	68%	57%
P	27	19.81	0.82	98%	89%	91%	82%	84%	89%	86%	86%	86%	86%
P	31	20.99	0.89	84%	68%	80%	59%	68%	66%	73%	68%	75%	73%
P	30	28.60	1.62	43%	59%	45%	52%	48%	43%	66%	55%	48%	57%

Chromatic thresholds measured at different background chromaticities using the CAD test paradigm are shown in Fig. 3. Variation in the ellipse parameters (other than the center point) was not statistically significant at the 5% level, including a minimal change in the ellipse orientation which varied between 63-70 degrees [within 95% CI for the mean]. When the ellipse parameters are converted to RG and YB CAD units the observed variability is within the established limits for within-subject variability on the CAD test. [26]

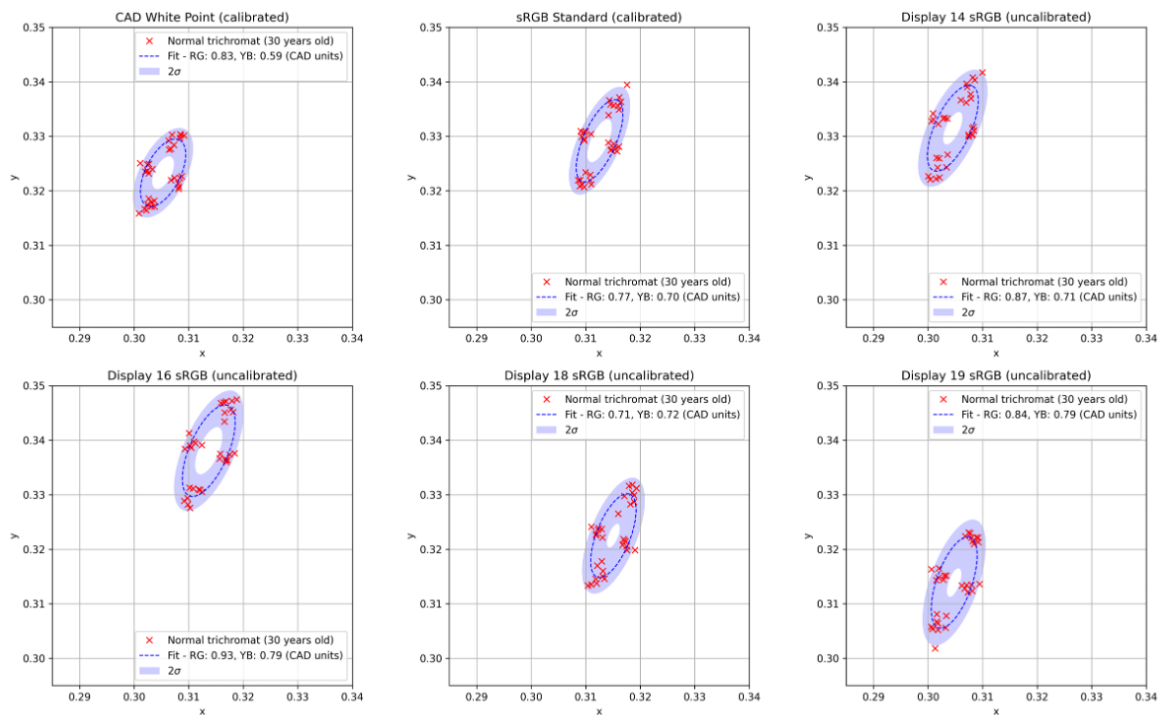


Fig. 3. Ellipses fitted to CAD thresholds measured in 32 directions in CIE x,y color space at 6 different background chromaticities. The 2σ variation is shown graphically by the shaded region and the equivalent CAD thresholds for each fitted ellipse are shown in each inset.

Discussion

The calibration of visual displays is the primary barrier and differentiator between robust, established measures of display-based methods and at-home color screening. How methods of color screening vary on sRGB visual displays in individuals with normal trichromacy and congenital color deficiency is therefore of significant interest for at-home testing. When carried out using sRGB visual displays the RG sensitivity [95% CI] of the CVS test is far lower (0.56 [0.46 – 0.64]) than the outcome when the test is carried out on spectrally calibrated visual displays (N=430, 1.00 [0.98 – 1.00]). [32] The specificity of the CVS test when participants are tested on uncalibrated sRGB visual displays is statistically equivalent to visually calibrated displays (sRGB: 1.00 [0.98 – 1.00]; calibrated: 0.99 [0.97 – 1.00]). The sensitivity does however vary markedly across different sRGB visual displays, ranging from

0.27 (0.01 - 0.57) to 0.73 (0.43 - 0.90), highlighting the differences one may expect across sRGB visual displays found in an at-home setting.

The variation between visual displays is not unexpected given the variability observed in their photo- and radiometric properties as shown in Fig. 1, Fig. 2 and Table 1. Less expected perhaps is the clear individual variation in CVS outcome observed in color deficient observers when assessed using sRGB visual displays (Table 2). Some color deficient participants (1 deuteranomalous trichromat, 1 deuteranope, and a protanope as classified by the Nagel anomaloscope) passed the CVS on sRGB visual displays, despite having 'moderate' or 'severe' color vision deficiency. [27] It should be noted that all of the participants shown in Table 2 failed the CVS (and were correctly identified as having RG loss) when the same CVS test was carried out using a spectrally calibrated visual display.

A source of variability across sRGB visual displays is the difference in the chromaticities of primary colors, as shown in Fig. 2. Results from an exploratory investigation into the impact of background chromaticity upon chromatic thresholds (shown in Fig. 3) revealed that variations observed for different background chromaticities fall within the expected mean, within subject variability of ~10% of the mean threshold. [26] Even at the same background chromaticity, the estimated parameters of each ellipse may vary by as much as 2σ . This suggests that shifts in background chromaticities are unlikely to account for the unexpected error patterns observed in subjects with congenital color deficiencies. Instead, subject-specific factors—such as variations in cone optical density, differences in L/M cone ratios, and residual luminance contrast signals—are more likely contributors to the observed outcome. However, as noted by Regan et al., the discrimination ellipses for CVDs, not normal trichromats, may rotate for different backgrounds [33] Further analyses with a larger cohort, including both normal trichromats and CVDs, is required to quantify fully the contribution of the different background chromaticities measured in this study.

Whilst variation in the maximum luminance of each display was observed, this alone may not explain the results, unless this variation also causes different changes in the gamma values of the primary colors of the display. We know that provided the chromaticity of the background remains relatively unchanged, RG thresholds remain stable down to ~ 2.5 cd/m². [34] Given the relatively high maximum luminances of the sRGB displays employed in this study (see Table 1) one can be confident that any differences in the mean background luminance amongst the sRGB displays cannot cause the poor sensitivity for detection of congenital RG deficiency observed in this study.

A limitation of the present work is that gamma values were only measured for a subset of displays, limiting the generalizability of these findings. Whilst visual display 19 had the

lowest sensitivity and differed most from the sRGB standard in terms of its white gamma, displays 12 and 13 were measured to have white gamma values matching the sRGB standard, yet both displays were found to have a low sensitivity (0.55). While deviations in gamma values may affect results for mildly color-deficient individuals, they fail, at least in isolation, to explain the significant variation in error responses among those with severe color deficiencies.

Another limitation is the lack of individuals with YB loss. Further work with a population that includes such individuals is required to establish the CVS's efficacy in screening for acquired color defects which result solely in YB loss. However, in cases of acquired loss where both RG and YB chromatic systems are impacted, the current study provides a partial picture of the CVS's screening accuracy. The consistency between the classification made by the CAD test and the Nagel anomaloscope further supports these findings.

An important finding that emerges from this study is the need to assess and test individuals with confirmed color deficiency if one wishes to validate a new method of color assessment using uncalibrated visual displays. The high specificity of the CVS on sRGB displays is in stark contrast to the reduced sensitivity, determined by anomaloscopy and CAD test classification. Modern validation studies often opt to use a screening test, such as the Ishihara or American Optical Hardy-Rand-Rittler pseudoisochromatic plates, as the reference measure. [8,10] The best outcome a test can achieve in a validation study is 100% agreement with the reference measure (or gold standard) used. As such, the choice to use any test that, depending upon the protocol used, has been reported to have a sensitivity as low as 0.80 or a specificity as low as 0.67, has clear limitations and cannot, nor should be, overlooked.

This work highlights the importance of full display calibration in addition to the limitations present in current methods of 'at-home' color screening. The reduction in sensitivity when one moves from calibrated to uncalibrated sRGB visual displays is significant and clinically relevant. Full display calibration is therefore required to improve the RG sensitivity of the CVS.

Conclusion

An investigation into at-home color vision screening was carried out to establish the types of light scatter present in visual display measurements, the efficacy of the CVS test on sRGB visual displays, and factors contributing to the variability observed in the CVS test outcome on such displays. The effect of 'dark light' within the display and the light scatter within the photometer head in particular should be minimized when calibrating visual displays. Whilst

within-display scatter contributes a small and consistent baseline level of 'dark light', the signal generated as a result of internal scattering of light within the photometer head increases with background size and luminance and can affect significantly the measured luminance of the area of interest during calibration.

In conclusion, the results suggest that full spectral and luminance calibration of visual displays are needed in order to improve the RG sensitivity of the CVS. The specificity of the test is statistically equivalent across sRGB visual displays and fully calibrated visual displays. Variations in maximum luminance values across sRGB visual displays do not fully explain the observed results, notably the within-participant variation. Different changes in the gamma functions of the primary colors can introduce significant errors in both the luminance and the chromaticity of the test stimuli and may account for the poor sensitivity of the CVS when used with uncalibrated displays. Further work is needed to establish whether changes in gamma functions account for poor RG sensitivity that can be achieved on uncalibrated displays. This work highlights the limitations of current at-home color screening methods and the necessity for improved calibration to enhance the accuracy of color vision assessments.

Funding. The COLT Foundation; UK Civil Aviation Authority.

Acknowledgments. We wish to thank and acknowledge the COLT Foundation and the UK Civil Aviation Authority for their financial support. We also wish to thank Stuart Mitchell who has supported us from the UK CAA.

Disclosures. The authors declare no conflicts of interest. Professor John Barbur designed all the Advanced Vision and Optometric Tests (AVOT), including the CAD test, but no longer has any commercial interests in these tests or current conflict of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

- 300 1. S. S. Z. Sii, C. S. Chean, H. Kuht, C. Bunce, M. G. Thomas, and S. R. Rufai, "Home-
301 based screening tools for amblyopia: a systematic review," *Eye (Basingstoke)* **37**,
302 2649–2658 (2023).
- 303 2. S. I. Gray, C. Bevan, S. Campbell, and K. Cater, "Space Vision: Developing a Game
304 for Vision Screening and Home-Based Monitoring," *Proc. ACM Hum.-Comput.*
305 *Interact.* **5**, (2021).
- 306 3. E. B. M. Elsman, S. Q. Lee, H. P. A. van der Aa, F. van Nassau, R. P. L. Wisse, O.
307 R. Maarsingh, J. E. E. Keunen, G. H. M. B. van Rens, and R. M. A. van Nispen, "The
308 evaluation of an online nurse-assisted eye-screening tool in older adults receiving
309 home healthcare," *Ophthalmic and Physiological Optics* **43**, 725–737 (2023).
- 310 4. M. D. Crossland, T. M. Dekker, J. Hancox, M. Lisi, T. A. Wemyss, and P. B. M.
311 Thomas, "Evaluation of a Home-Printable Vision Screening Test for Telemedicine,"
312 *JAMA Ophthalmol* **139**, 271–277 (2021).
- 313 5. E. Y. Chew, T. E. Clemons, S. B. Bressler, M. J. Elman, R. P. Danis, A. Domalpally,
314 J. S. Heier, J. E. Kim, and R. A. Garfinkel, "Randomized trial of the ForeseeHome
315 monitoring device for early detection of neovascular age-related macular
316 degeneration. The HOme Monitoring of the Eye (HOME) study design - HOME Study
317 report number 1," *Contemp Clin Trials* **37**, 294–300 (2014).
- 318 6. X. Han, J. Scheetz, S. Keel, C. Liao, C. Liu, Y. Jiang, A. Müller, W. Meng, and M. He,
319 "Development and validation of a smartphone-based visual acuity test (Vision at
320 home)," *Transl Vis Sci Technol* **8**, (2019).
- 321 7. T. Tang, L. Álvaro, J. Alvarez, J. Maule, A. Skelton, A. Franklin, and J. Bosten,
322 "ColourSpot, a novel gamified tablet-based test for accurate diagnosis of color vision
323 deficiency in young children," *Behav Res Methods* **54**, 1148–1160 (2022).
- 324 8. V. Karampatakis, D. Almaliotis, L. Karamitopoulos, G. Kalliris, and S. Almpandou, "A
325 Novel Smartphone-Based Color Test for Detection of Color Vision Defects in Age
326 Related Macular Degeneration," *J Ophthalmol* **2022**, 14–16 (2022).
- 327 9. J. Seshadri, J. Christensen, V. Lakshminarayanan, and C. J. Bassi, "Evaluation of the
328 new web-based "Colour Assessment and Diagnosis" test," *Optometry and Vision*
329 *Science* **82**, 882–885 (2005).

- 330 10. D. de Fez, M. J. Luque, L. Matea, D. P. Piñero, and V. J. Camps, "New iPad-based
331 test for the detection of color vision deficiencies," *Graefe's Archive for Clinical and*
332 *Experimental Ophthalmology* **256**, 2349–2360 (2018).
- 333 11. L. C. Nguyen, E. Y.-L. Do, A. Chia, Y. Wang, and H. B.-L. Duh, "DoDo game, a color
334 vision deficiency screening test for young children," in *Proceedings of the SIGCHI*
335 *Conference on Human Factors in Computing Systems*, CHI '14 (Association for
336 Computing Machinery, 2014), pp. 2289–2292.
- 337 12. L. Nguyen, W. Lu, E. Y.-L. Do, A. Chia, and Y. Wang, "Using digital game as clinical
338 screening test to detect color deficiency in young children," in *Proceedings of the*
339 *2014 Conference on Interaction Design and Children*, IDC '14 (Association for
340 Computing Machinery, 2014), pp. 337–340.
- 341 13. W. D. Thomson and B. Evans, "A new approach to vision screening in schools,"
342 *Ophthalmic and Physiological Optics* **19**, 196–209 (1999).
- 343 14. K. R. Gegenfurtner and D. C. Kiper, "Color Vision," *Annu Rev Neurosci* **26**, 181–206
344 (2003).
- 345 15. M. P. Simunovic, "Acquired color vision deficiency," *Surv Ophthalmol* **61**, 132–155
346 (2016).
- 347 16. M. Pacheco-Cutillas, D. F. Edgar, and A. Sahraie, "Acquired colour vision defects in
348 glaucoma---their detection and clinical significance," *British Journal of Ophthalmology*
349 **83**, 1396–1402 (1999).
- 350 17. T. H. Margrain, J. Birch, and C. G. Owen, "Colour vision requirements of firefighters,"
351 *Occup Med (Chic Ill)* **46**, 114–124 (1996).
- 352 18. A. Stockman, *Report on Colour Vision Standards for STU Officers* (2018).
- 353 19. T. Carter and J. Barbur, *Colour Vision Assessment for Maritime Navigational*
354 *Lookout : Review for UK Maritime and Coastguard Agency (MCA)* (2015).
- 355 20. H. Kolb, "How the Retina Works," *Am Sci* **91**, 28 (2003).
- 356 21. A. Stockman, D. I. A. MacLeod, and N. E. Johnson, "Spectral sensitivities of the
357 human cones," *Journal of the Optical Society of America A* **10**, 2491 (1993).
- 358 22. F. G. Rauscher, C. M. Chisholm, D. F. Edgar, and J. L. Barbur, "Assessment of novel
359 binocular colour, motion and contrast tests in glaucoma," *Cell Tissue Res* **353**, 297–
360 310 (2013).

- 361 23. M. O'Neill-Biba, S. Sivaprasad, M. Rodriguez-Carmona, J. E. Wolf, and J. L. Barbur,
362 "Loss of chromatic sensitivity in AMD and diabetes: A comparative study," *Ophthalmic*
363 *and Physiological Optics* **30**, 705–716 (2010).
- 364 24. S. J. Dain, B. Kwan, and L. Wong, "Consistency of color representation in smart
365 phones," *J. Opt. Soc. Am. A* **33**, A300–A305 (2016).
- 366 25. International Electrotechnical Commission (IEC), *Multimedia Systems and Equipment*
367 *- Colour Measurement and Management - Part 2-1: Colour Management - Default*
368 *RGB Colour Space - sRGB (IEC 61966-2-1:1999)* (1999).
- 369 26. J. L. Barbur, M. Rodriguez-Carmona, and B. E. W. Evans, "Color vision assessment-
370 3. An efficient, two-step, color assessment protocol," *Color Res Appl* **46**, 33–45
371 (2021).
- 372 27. J. L. Barbur and M. Rodriguez-Carmona, "Colour vision requirements in visually
373 demanding occupations," *Br Med Bull* **122**, 51–77 (2017).
- 374 28. A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE*
375 *Trans Pattern Anal Mach Intell* **21**, 476–480 (1999).
- 376 29. R. Halir and J. Flusser, "Numerically stable direct least squares fitting of ellipses," in
377 *Proc. 6th International Conference in Central Europe on Computer Graphics and*
378 *Visualization* (IEEE, 1998), pp. 125–132.
- 379 30. P. Zsombor-Murray, "Direct and specific least-square fitting of hyperbolæ and
380 ellipses," *J Electron Imaging* **13**, 492 (2004).
- 381 31. A. Affan, "Ellipse-Fit-Uncertainty," .
- 382 32. B. Evans, M. Rodriguez Carmona, F. Rauscher, E. Llapashtica, V. Koefoed, F.
383 Ziemssen, R. Nitsche, A. Farini, E. Baldanzi, L. Robledo, A. Douglass, M. Baker, R.
384 Quast, S. Roelcke, S. Ho, and J. Barbur, "Multi-centre evaluation of the Colour Vision
385 Screener test," *Journal of the Optical Society of America A* (2025).
- 386 33. B. C. Regan, J. P. Reffin, and J. D. Mollon, "Luminance noise and the rapid
387 determination of discrimination ellipses in colour deficiency," *Vision Res* **34**, 1279–
388 1299 (1994).
- 389 34. J. L. Barbur, E. Llapashtica, A. T. Rider, and A. Stockman, "Photopic, Mesopic and
390 Scotopic Vision and Changes in Visual Performance," in *Reference Module in*
391 *Neuroscience and Biobehavioral Psychology*, D. A. Dartt, ed. (Elsevier, 2024), pp.
392 323–331.

393

394

395

396

397

398